

From: Myers, Carla
Sent: Wednesday, July 13, 2005 7:06 AM
To: STIC-Biotech/ChemLib
Subject: sequence search for 08/974,584

(STIC)

JUL 13 2005

RECEIVED

Please search commercial and interference databases for:

- 1) nucleic acids encoding SEQ ID NO: 118 (1132 amino acids);
- 2) proteins comprising SEQ ID NO: 118 (1132 amino acids)
- 3) nucleic acids encoding each of SEQ ID NO: 139, 143, 144, 146, 147, and 16
- 4) nucleic acids encoding each of SEQ ID NO: 139, 143, 144, 146, 147 and 17

** if it is not possible to search 3 and 4 (i.e., nucleic acids encoding multiple SEQ ID NOs), then please perform a separate search for nucleic acids encoding SEQ ID NO: 139, 143, 144, 146, 147, 16 and 17.

SEQ 139=11 amino acids; SEQ 143=5 amino acids; SEQ 144=22 amino acids; SEQ 147=13 amino acids;
SEQ 16=50 amino acids; SEQ 17=43 amino acids

The CRF has been entered: http://expoweb1:8001/cgi-bin/expo/BioInfo/bioquery.pl?APPL_ID=08974584

**Please provide a printout of the first 40 results.

Thank you-

Carla Myers
AU 1634
Remsen Bldg / Rm 2E79
Mailbox: REM 2C70
571-272-0747

61/9 03p/17 03h/13 04/36 05p/4 05h/20
06p/10 06h/11

* myers 08974584*
* us08974584*

STAFF USE ONLY

Searcher: PCB
Searcher Phone: 2-
Date Searcher Picked up: 7-27-05
Date Completed: 7-27-05
Searcher Prep/Rev. Time: 20
Online Time: 18

Type of Search

NA#: 3 AA#: 1
Interference: SPDI:
S/L: Oligomer:
Encode/Transi:
Structure#: Text:
Inventor: Litigation:

Vendors and cost where applicable

STN: 06h, IG
DIALOG: IG
QUESTEL/ORBIT: IG
LEXIS/NEXIS: IG
SEQUENCE SYSTEM: IG
WWW/Internet: IG
Other(Specify): IG

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:18:34 ; Search time 4179.29 Seconds
(without alignments)
938.107 Million cell updates/sec

Title: SEQ139-143-144-146-147-17
Perfect score: 233
Sequence: 1 XXXRXXPKXXXRXIXXXXX.....WXTXXXXXXXRXRXXXW 103

Scoring table: BLOSUM62
-Q=/cg2.1/USPTO spo01/MYERS08974584/runat_25072005_102706_5075/app_query.fasta_1.526
-DB=EST_QFMT=fastep -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=MYERS08974584 @CGN 1.1 3556 @runat_25072005_102706_5075 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSBLOCK=100 -JONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cg2.1/USPTO spo01/MYERS08974584/runat_25072005_102706_5075/app_query.fasta_1.526
-DB=EST_QFMT=fastep -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=MYERS08974584 @CGN 1.1 3556 @runat_25072005_102706_5075 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSBLOCK=100 -JONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gest1:*
9: gb_gest2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	39	16.7	352	BP427888	BP427888 BP427888
C 2	39	16.7	437	BUS06243	BUS06243 AGENCOURT
C 3	39	16.7	576	BP760208	BP760208 BP760208
C 4	39	16.7	601	BP766928	BP766928 BP766928
C 5	39	16.7	642	BP771659	BP771659 BP771659
C 6	38	16.3	1047	BQ224762	BQ224762 AGENCOURT
C 7	37	15.9	939	CL398497	CL398497 ZMMBBB039
C 8	36.8	15.8	851	BUS09508	BUS09508 AGENCOURT
C 9	36.4	15.6	1058	BH725011	BH725011 BOMNB89TF

10	36.1	15.5	181	9	CG784907	CG784907 RRR157 Ba
11	36.1	15.5	197	6	CG887155	CG887155 RRS688 Ba
12	36.1	15.5	216	6	CB178950	CB178950 in99504.y
13	36.1	15.5	357	9	CG784906	CG784906 RRR158 Ba
14	36.1	15.5	408	7	CK225321	CK225321 703207690
15	36.1	15.5	563	7	CN722338	CN722338 E0841A10-
16	36.1	15.5	578	7	CN717076	CN717076 E0737G08-
17	36.1	15.5	624	7	CN699623	CN699623 E0423H10-
18	36.1	15.5	665	2	BB616800	BB616800 BB616800
19	36.1	15.5	741	7	CF733575	CF733575 UI-M-HB0-
20	36.1	15.5	746	7	CK780407	CK780407 UI-M-GV0-
21	36.1	15.5	748	6	CD803163	CD803163 UI-M-GV0-
22	36.1	15.5	768	7	CO433173	CO433173 UI-M-HX0-
23	36.1	15.5	859	7	CO807249	CO807249 AGENCOURT
24	36.1	15.5	3232	3	AK030089	AK030089 Mus muscu
25	36.1	15.5	4003	3	AK028324	AK028324 Mus muscu
26	36.1	15.5	4024	3	AK030951	AK030951 Mus muscu
27	36.1	15.5	4306	3	AK032405	AK032405 Mus muscu
28	36.1	15.5	4942	3	AK031688	AK031688 Mus muscu
29	35.9	15.4	548	7	W43412	W43412 22805 CD4-1
30	35.9	15.4	598	5	BP561807	BP561807 BP561807
31	35.9	15.4	651	3	CNS0A2QM	BP829258 Arabidops
32	35.9	15.4	680	5	BX836890	BX836890 BX836890
33	35.9	15.4	681	1	AV822745	AV822745 AV822745
34	35.9	15.4	692	3	CNS0A2P9	BX829061 Arabidops
35	35.9	15.4	693	3	CNS09YGP	BX841609 Arabidops
36	35.9	15.4	753	8	AZ869871	AZ869871 2M0182H11
37	35.9	15.4	758	3	CNS0A2LE	BX828731 Arabidops
38	35.9	15.4	893	7	CF652848	CF652848 80-L02052
39	35.9	15.4	900	7	CV242818	CV242818 WS02516.B
40	35.9	15.4	928	9	CL514461	CL514461 SAIL_889
41	35.3	15.2	563	9	BX989710	BX989710 Forward s
42	35.1	15.1	582	1	AJ799564	AJ799564 AJ799564
43	35.1	15.1	716	1	AJ560032	AJ560032 AJ560032
44	35	15.0	610	5	BP154492	BP154492 BP154492
45	35	15.0	669	8	AQ840198	AQ840198 hdx0052M
46	35	15.0	794	7	CK305937	CK305937 SB02031B2
47	35	15.0	1107	5	BQ618909	BQ618909 RNOSQ1C0
48	35	15.0	1107	5	BQ619092	BQ619092 RNOSQ1C0
49	35	15.0	1107	5	BQ619227	BQ619227 RNOSQ5B1
50	35	15.0	1107	5	BQ619307	BQ619307 RNOSQ6B0
51	35	15.0	1107	5	BQ619325	BQ619325 RNOSQ6B0
52	34.9	15.0	449	8	AZ216917	AZ216917 Sheared D
53	34.9	15.0	554	9	TA271D10Q	AL488262 T. brucei
54	34.4	14.8	416	5	BY224521	BY224521 BY224521
55	34.2	14.7	648	9	CE461221	CE461221 tigr-g88-
56	34.1	14.6	206	8	BZ660880	BZ660880 SALK_0243
57	34	14.6	365	2	BF395855	BF395855 UI-R-BS2-
58	34	14.6	424	8	AQ084051	AQ084051 HS_2230.B
59	34	14.6	429	2	BF409184	BF409184 UI-R-BT1-
60	34	14.6	431	1	AA045746	AA045746 zk68b09.r
61	34	14.6	450	8	AZ129520	AZ129520 OSJNB010
62	34	14.6	499	6	CF057037	CF057037 QCO27a12.
63	34	14.6	519	2	BF853241	BF853241 MR2-EN009
64	34	14.6	601	8	BH473238	BH473238 BOGYJ22TF
65	34	14.6	607	2	BF398120	BF398120 UI-R-BS2-
66	34	14.6	625	7	CK004457	CK004457 AGENCOURT
67	34	14.6	750	8	AZ656688	AZ656688 1M0532H20
68	34	14.6	751	8	BZ612798	BZ612798 WHAAS17TF
69	33.9	14.6	847	5	BUI38183	BUI38183 603134394
70	33.9	14.5	353	8	AQ927217	AQ927217 RPCI-23-2
71	33.9	14.5	445	7	CN561085	CN561085 taf78b11.
72	33.8	14.5	458	5	BQ391944	BQ391944 NISC mq21
73	33.8	14.5	652	7	CN100249	CN100249 EC2CA17B
74	33.8	14.5	685	4	BG369595	BG369595 HVSMB1002
75	33.2	14.2	1029	9	CL035455	CL035455 CH216-39N
76	33.1	14.2	348	7	T90943	T90943 yd57d12.e1
77	33.1	14.2	385	1	AL881564	AL881564 AL881564
78	33.1	14.2	620	2	AK306948	AK306948 ef50q09.y
79	33	14.2	100	2	BF840692	BF840692 IL0-HT106
80	33	14.2	132	2	BB807052	BB807052 BB807052
81	33	14.2	152	1	AJ346688	AJ346688 sHR-00001
82	33	14.2	155	4	BG378472	BG378472 UI-R-CU0-

83	33	14.2	156	1	AL797816	AL797816
84	33	14.2	165	1	AV157928	AV157928
85	33	14.2	171	1	AL883683	AL883683
86	33	14.2	173	2	BE089816	RC5-BT070
87	33	14.2	176	2	BF79876	RC2-ET018
88	33	14.2	184	8	AA338728	AA069J21
89	33	14.2	191	2	BF118430	1U095D09.X
90	33	14.2	197	1	AI609913	AI63H08.X
91	33	14.2	198	7	CF209635	CF209635
92	33	14.2	211	6	CB189535	CB189535
93	33	14.2	216	1	AL868343	AL868343
94	33	14.2	220	1	AA25807	AA25807
95	33	14.2	229	7	CF735736	CF735736
96	33	14.2	233	2	BB132151	BB132151
97	33	14.2	240	2	BE117386	UI-R-B81-
98	33	14.2	243	7	CV321951	CM3-HT096
99	33	14.2	246	5	BK629569	BK629569
100	33	14.2	246	7	CV323271	CM4-AN007

ALIGNMENTS

RESULT 1	
BP427888/c	
LOCUS	352 bp mRNA linear EST 24-MAY-2004
DEFINITION	Mus musculus cerebellum E18-P56 Mus musculus cDNA clone
ACCESSION	BP427888
VERSION	FD1633. mRNA sequence.
KEYWORDS	BP427888.1 GI:47589962
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
JOURNAL	1 (bases 1 to 352)
COMMENT	Furuichi,T. and Sato,A. Gene expression during the postnatal cerebellar development Unpublished (2004) Contact: Teiichi Furuichi Laboratory for Molecular Neurogenesis RIKEN Brain Science Institute 2-1 Hirotsawa, Wako, Saitama 351-0198, Japan Tel: 81-48-467-5906 Fax: 81-48-467-6079 Email: mol-neurogenisis@brain.riken.go.jp, URL: http://www.brain.riken.go.jp/labs/lmu/index.html .

```

ORIGIN
Alignment Scores:
Pred. No.: 164
Score: 39.00
Percent Similarity: 31.58%
Best Local Similarity: 31.58%
Query Match: 16.74%
DB: 5
Length: 352
Matches: 6
Conservative: 0
Mismatch: 13
Indels: 0
Gaps: 0
SEQ139-143-144-146-147-17 (1-103) x BP427888 (1-352)
QY 69 Trp*****phoPhoTrp***ThrGlu 87
|||
103 TGGAGTCCAGATCCATCGACGCAAGTTCTCAGTTCCTCTTTCTGGGAAACAGAA 47
Db
RESULT 2

```

```

BU506243/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
    source
        BU506243
        AGENCOURT 10017055 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494302
        5', mRNA Sequence.
        BU506243
        BU506243.1 GI:22812476
        Mus musculus (house mouse)
        Mus musculus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
        1 (bases 1 to 437)
        NIH-MGC http://img.nci.nih.gov/.
        National Institutes of Health, Mammalian Gene Collection (MGC)
        Unpublished (1999)
        Contact: Robert Strausberg, Ph.D.
        Email: cgabbs@mail.nih.gov
        Tissue Procurement: The Cepko Laboratory
        cDNA Library Preparation: Life Technologies, Inc.
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Agencourt Bioscience Corporation
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM14050 row: e column: 23
        High quality sequence stop: 436.
        Location/Qualifiers
            1..437
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:6494302"
                /tissue_type="retina"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 94"
                /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally; oligo-dt primed.
                Average insert size 3.3 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH MGC library."

```

Note: this is a non-MOC library.

ORIGIN

Alignment Scores:

Pred. No.:	207	Length:	437
Score:	39.00	Matches:	6
Percent Similarity:	31.58%	Conservative:	0
Best Local Similarity:	31.58%	Mismatches:	13
Query Match:	16.74%	Indels:	0
DB:	5	Gaps:	0

SEQ139-143-144-146-147-17 (1-103) x BU506243 (1-437)

Qy	69	Tp		*****PhePheTrp***ThrGlu	87
Db	138	TGGAGTCCAGATCCATCGCAGCAAGTTCTCAGTTCCTCTTTTCGGGAACAGAA			82

RESULT 3

BP760208

LOCUS

DEFINITION

BP760208 mouse (C57BL/6) pancreatic islet library with
recombination-based method Mus musculus cDNA clone mib31028 3',
mRNA sequence.

ACCESSION

BP760208

VERSION

BP760208.1 GI:50218906

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.

REFERENCE

1 (bases 1 to 576)

AUTHORS

Nishimura,M., Yokoi,N., Miki,T., Horikawa,Y., Yoshioka,H.,
Takeda,J., Ohara,O. and Seino,S.

TITLE

Construction of a multi-functional cDNA library specific for mouse

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 00:55:25 ; Search time 1188.27 Seconds
(without alignments)
560.502 Million cell updates/sec

Title: SEQ139-143-144-146-147-17
Perfect score: 233
Sequence: 1 XXXRXXPPXXXXXXIXXXXX.....WXTXXXXXXXIXXXXXX 103

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.1
Ygapop 10.0, Ygapext 0.1
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Total number of hits satisfying chosen parameters: 14555652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/MYERS08974584/runat_25072005_102708_5180/app.query.fasta_1.526
-DB=Published Applications NA -QWMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MAIRIX=BLOSUM62
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000
-USER=MYERS08974584 @CGN 1 1.456 @runat_25072005_102708_5180 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	35.9	15.4	368	US-09-770-791-397	Sequence 397, App
2	35.9	15.4	434	US-09-924-035A-656	Sequence 656, App
3	35.9	15.4	639	US-09-938-842A-2291	Sequence 2291, App
4	35.9	15.4	639	US-09-938-842A-2291	Sequence 2291, App
5	35.9	15.4	2321	US-10-739-933-1097	Sequence 1097, App
6	35	15.0	455	US-10-437-963-9511	Sequence 9511, App
7	35	15.0	600	US-10-972-079-62547	Sequence 62547, A
8	34.1	14.6	784	US-10-225-066A-655	Sequence 655, App
9	34.1	14.6	784	US-10-374-780A-2237	Sequence 2237, App
10	34	14.6	404	US-10-424-599-40084	Sequence 40084, A
11	33	14.2	102	US-10-472-928-1683	Sequence 1683, App
12	33	14.2	140	US-09-783-590-278	Sequence 278, App
13	33	14.2	193	US-10-437-963-102089	Sequence 102089,
14	33	14.2	201	US-10-719-993-25917	Sequence 25917, A
15	33	14.2	282	US-10-424-599-57538	Sequence 57538, A
16	33	14.2	326	US-10-767-795-4667	Sequence 4667, App
17	33	14.2	330	US-09-783-590-11246	Sequence 11246, A
18	33	14.2	376	US-09-880-107-1031	Sequence 1031, App
19	33	14.2	460	US-10-425-115-80762	Sequence 80762, A
20	33	14.2	488	US-10-060-036-1633	Sequence 1633, App
21	33	14.2	521	US-10-425-115-36983	Sequence 36983, A
22	33	14.2	529	US-10-767-701-5515	Sequence 5515, App
23	33	14.2	539	US-10-424-599-29238	Sequence 29238, A
24	33	14.2	572	US-10-021-323-12142	Sequence 12142, A
25	33	14.2	593	US-10-027-632-271755	Sequence 271755,
26	33	14.2	593	US-10-027-632-271755	Sequence 271755,
27	33	14.2	595	US-10-767-795-4666	Sequence 4666, App
28	33	14.2	616	US-10-027-632-265805	Sequence 265805,
29	33	14.2	616	US-10-027-632-265806	Sequence 265806,
30	33	14.2	616	US-10-027-632-265805	Sequence 265805,
31	33	14.2	616	US-10-027-632-265806	Sequence 265806,
32	33	14.2	619	US-10-027-632-196891	Sequence 196891,
33	33	14.2	619	US-10-027-632-196892	Sequence 196892,
34	33	14.2	619	US-10-027-632-196893	Sequence 196893,
35	33	14.2	619	US-10-027-632-196891	Sequence 196891,
36	33	14.2	619	US-10-027-632-196892	Sequence 196892,
37	33	14.2	619	US-10-027-632-196893	Sequence 196893,
38	33	14.2	649	US-10-335-053-194	Sequence 194, App
39	33	14.2	652	US-10-357-930-52657	Sequence 52657, A
40	33	14.2	664	US-10-425-114-16192	Sequence 16192, A
41	33	14.2	667	US-10-027-632-287415	Sequence 287415,
42	33	14.2	667	US-10-027-632-287416	Sequence 287416,
43	33	14.2	667	US-10-027-632-287415	Sequence 287415,
44	33	14.2	667	US-10-027-632-287416	Sequence 287416,
45	33	14.2	721	US-10-027-632-260266	Sequence 260266,
46	33	14.2	721	US-10-027-632-260266	Sequence 260266,
47	33	14.2	889	US-10-363-345A-28829	Sequence 28829, A
48	33	14.2	889	US-10-363-345A-28830	Sequence 28830, A
49	33	14.2	889	US-10-363-345A-28829	Sequence 28829, A
50	33	14.2	889	US-10-363-345A-28830	Sequence 28830, A
51	33	14.2	900	US-10-425-115-160420	Sequence 160420,
52	33	14.2	987	US-10-424-599-90405	Sequence 90405, A
53	33	14.2	1113	US-10-474-776-76	Sequence 76, Appl
54	33	14.2	1162	US-10-956-157-1455	Sequence 1455, App
55	33	14.2	1162	US-10-956-157-6690	Sequence 6690, App
56	33	14.2	1171	US-10-369-493-36598	Sequence 36598, A
57	33	14.2	1171	US-10-363-345A-40425	Sequence 40425, A
58	33	14.2	1174	US-10-363-345A-40426	Sequence 40426, A
59	33	14.2	1174	US-10-363-345A-40425	Sequence 40425, A
60	33	14.2	1174	US-10-363-345A-40426	Sequence 40426, A
61	33	14.2	1201	US-10-425-115-89303	Sequence 89303, A
62	33	14.2	1312	US-10-425-114-25327	Sequence 25327, A
63	33	14.2	1383	US-10-437-963-100022	Sequence 100022,
64	33	14.2	1404	US-10-425-115-49114	Sequence 49114, A

```
65 33 14.2 1554 21 US-10-470-048B-537
66 33 14.2 1575 22 US-10-724-972A-3346
67 33 14.2 1631 13 US-10-027-632-258931
68 33 14.2 1631 13 US-10-027-632-258932
69 33 14.2 1631 13 US-10-027-632-258933
70 33 14.2 1631 13 US-10-027-632-258934
71 33 14.2 1631 17 US-10-027-632-258931
72 33 14.2 1631 17 US-10-027-632-258932
73 33 14.2 1631 17 US-10-027-632-258933
74 33 14.2 1631 17 US-10-027-632-258934
75 33 14.2 1988 13 US-10-027-632-97163
76 33 14.2 1988 13 US-10-027-632-97164
77 33 14.2 1988 13 US-10-027-632-97165
78 33 14.2 1988 17 US-10-027-632-97163
79 33 14.2 1988 17 US-10-027-632-97164
80 33 14.2 1988 17 US-10-027-632-97165
81 33 14.2 2418 9 US-09-938-842A-1478
82 33 14.2 2418 11 US-09-938-842A-1478
83 33 14.2 2563 9 US-09-782-051-1
84 33 14.2 2655 21 US-10-876-285-3
85 33 14.2 2760 20 US-10-739-930-157
86 33 14.2 3634 20 US-10-357-930-23184
87 33 14.2 3634 20 US-10-357-930-29050
88 33 14.2 3986 13 US-10-092-925-1
89 33 14.2 4105 17 US-10-062-674-1991
90 33 14.2 4333 20 US-10-357-930-21849
91 33 14.2 4333 20 US-10-357-930-22736
92 33 14.2 4333 20 US-10-357-930-27699
93 33 14.2 4333 20 US-10-357-930-28584
94 33 14.2 5201 19 US-10-717-597-213
95 33 14.2 5384 15 US-10-311-455-181
96 33 14.2 6915 15 US-10-311-455-379
97 33 14.2 18627 8 US-08-961-527-113
98 33 14.2 18627 17 US-10-158-844-113
99 33 14.2 44096 21 US-10-876-285-1
100 33 14.2 88853 13 US-10-087-192-85

ALIGNMENTS

RESULT 1
US-09-770-791-397
; Sequence 397, Application US/09770791
; Patent No. US20020062014A1
; GENERAL INFORMATION:
; APPLICANT: Grlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Gricker, Carlos A.
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2029 (PARA-018PRV)
; CURRENT APPLICATION NUMBER: US/09/770,791
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,480
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0

Sequence 537, App
Sequence 3346, Ap
Sequence 258931,
Sequence 258932,
Sequence 258933,
Sequence 258934,
Sequence 258931,
Sequence 258932,
Sequence 258933,
Sequence 258934,
Sequence 97163, A
Sequence 97164, A
Sequence 97165, A
Sequence 1478, Ap
Sequence 1478, Ap
Sequence 1, Appli
Sequence 3, Appli
Sequence 157, App
Sequence 23184, A
Sequence 29050, A
Sequence 1, Appli
Sequence 1991, Ap
Sequence 21849, A
Sequence 22736, A
Sequence 27699, A
Sequence 28584, A
Sequence 213, App
Sequence 181, App
Sequence 379, App
Sequence 113, App
Sequence 113, App
Sequence 1, Appli
Sequence 85, Appli

; SEQ ID NO 397
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(368)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-791-397
Alignment Scores:
Pred. No.: 183
Score: 35.90
Percent Similarity: 26.92%
Best Local Similarity: 26.92%
Query Match: 15.41%
Indels: 11
Gaps: 1
Length: 368
Matches: 7
Conservative: 0
Mismatch: 8
Indels: 11
Gaps: 1

SEQ139-143-144-146-147-17 (1-103) x US-09-770-791-397 (1-368)
Qy 62 AspAsp***Leu*****Trp***** 81
Db 231 GACGACTCTCTATCTGCCACGTGG-----CGG 257
Qy 82 PhePheTrp***ThrGlu 87
Db 258 CTCTTTGGACGACGGAG 275

RESULT 2
US-09-924-035A-656
; Sequence 656, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grlach, Jrn
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 656
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(434)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-656
Alignment Scores:
Pred. No.: 216
Score: 35.90
Percent Similarity: 26.92%
Best Local Similarity: 26.92%
Query Match: 15.41%
Indels: 9
Gaps: 1
Length: 434
Matches: 7
Conservative: 0
Mismatch: 8
Indels: 11
Gaps: 1

SEQ139-143-144-146-147-17 (1-103) x US-09-924-035A-656 (1-434)
Qy 62 AspAsp***Leu*****Trp***** 81
Db 293 GACGACTCTCTATCTGCCACGTGG-----CGG 319
Qy 82 PhePheTrp***ThrGlu 87
Db 320 CTCTTTGGACGACGGAG 337

RESULT 3
US-09-938-842A-2291
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:19:19 ; Search time 1974.08 Seconds
(without alignments)
85.375 Million cell updates/sec

Title: SEQ139-143-144-146-147-17

Perfect score: 233

Sequence: 1 XXXRXXPOXXXXXIXXXXXX.....WXTXXXXXXXXXXXXXXW 103

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.1
Ygapop 10.0, Ygapext 0.1
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/MYERS08974584/runat_25072005_102706_5094/app.query.fasta_1.526
-DB=Issued Patents NA -QWMT=fastap -SUFFIX=rnl -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER-MYERS08974584 @CNC 1.1 77 @runat_25072005_102706_5094 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	33	14.2	601	4	US-09-949-016-68055
C 2	33	14.2	601	4	US-09-949-016-72934
C 3	33	14.2	601	4	US-09-949-016-72935
C 4	33	14.2	601	4	US-09-949-016-206055
C 5	33	14.2	601	4	US-09-949-016-206056
C 6	33	14.2	601	4	US-09-949-016-206092
C 7	33	14.2	601	4	US-09-949-016-206093
C 8	33	14.2	1375	3	US-09-134-003C-2620
C 9	33	14.2	2226	2	US-08-896-003-2
C 10	33	14.2	2226	3	US-09-347-878-3
C 11	33	14.2	18627	3	US-08-961-527-113
C 12	33	14.2	104428	4	US-09-949-016-12737

Sequence 13814, A	US-09-949-016-13814	4	14.2	104429	33	13	C 13
Sequence 17561, A	US-09-949-016-17561	4	14.2	113379	33	14	C 14
Sequence 17562, A	US-09-949-016-17562	4	14.2	113379	33	15	C 15
Sequence 17202, A	US-09-949-016-17202	4	14.2	189560	33	16	C 16
Sequence 11869, A	US-09-949-016-11869	4	14.2	194714	33	17	C 17
Sequence 13875, A	US-09-949-016-13875	4	14.2	221545	33	18	C 18
Sequence 15725, A	US-09-949-016-15725	4	14.2	264358	33	19	C 19
Sequence 13747, A	US-09-949-016-13747	4	14.2	264665	33	20	C 20
Sequence 21066, A	US-09-513-999C-21066	3	14.1	458	32.8	21	C 21
Sequence 391, App	US-09-161-241-3	3	14.1	801	32.8	22	C 22
Sequence 1085, App	US-09-919-039-391	4	14.1	1215	32.8	23	C 23
Sequence 23306, A	US-09-976-594-1085	4	14.1	1797	32.8	24	C 24
Sequence 48371, A	US-09-949-016-48371	4	13.8	48119	32.2	26	C 26
Sequence 12003, A	US-09-949-016-12003	4	13.8	48119	32.2	27	C 27
Sequence 13177, A	US-09-018-584A-23	3	13.7	359	32	29	C 29
Sequence 23, Appl	US-09-784-423-23	3	13.7	359	32	30	C 30
Sequence 23, Appl	US-08-313-075A-55	1	13.7	386	32	31	C 31
Sequence 55, Appl	US-09-976-594-13	4	13.7	460	32	32	C 32
Sequence 13, Appl	US-09-270-767-11994	4	13.7	462	32	33	C 33
Sequence 206, App	US-09-107-532A-206	4	13.7	468	32	34	C 34
Sequence 444, App	US-10-101-464A-444	4	13.7	516	32	35	C 35
Sequence 10088, A	US-09-621-976-10088	4	13.7	519	32	36	C 36
Sequence 2892, A	US-09-949-016-2892	4	13.7	601	32	37	C 37
Sequence 28910, A	US-09-949-016-28910	4	13.7	601	32	38	C 38
Sequence 31502, A	US-09-949-016-31502	4	13.7	601	32	39	C 39
Sequence 31503, A	US-09-949-016-31503	4	13.7	601	32	40	C 40
Sequence 32526, A	US-09-949-016-32526	4	13.7	601	32	41	C 41
Sequence 46140, A	US-09-949-016-46140	4	13.7	601	32	42	C 42
Sequence 46141, A	US-09-949-016-46141	4	13.7	601	32	43	C 43
Sequence 61322, A	US-09-949-016-61322	4	13.7	601	32	44	C 44
Sequence 68082, A	US-09-949-016-68082	4	13.7	601	32	45	C 45
Sequence 78126, A	US-09-949-016-78126	4	13.7	601	32	46	C 46
Sequence 79760, A	US-09-949-016-79760	4	13.7	601	32	47	C 47
Sequence 125957, A	US-09-949-016-125957	4	13.7	601	32	48	C 48
Sequence 137324, A	US-09-949-016-137324	4	13.7	601	32	49	C 49
Sequence 145882, A	US-09-949-016-145882	4	13.7	601	32	50	C 50
Sequence 145950, A	US-09-949-016-145950	4	13.7	601	32	51	C 51
Sequence 146218, A	US-09-949-016-146218	4	13.7	601	32	52	C 52
Sequence 152985, A	US-09-949-016-152985	4	13.7	601	32	53	C 53
Sequence 152986, A	US-09-949-016-152986	4	13.7	601	32	54	C 54
Sequence 160719, A	US-09-949-016-160719	4	13.7	601	32	55	C 55
Sequence 160720, A	US-09-949-016-160720	4	13.7	601	32	56	C 56
Sequence 160721, A	US-09-949-016-160721	4	13.7	601	32	57	C 57
Sequence 163504, A	US-09-949-016-163504	4	13.7	601	32	58	C 58
Sequence 163505, A	US-09-949-016-163505	4	13.7	601	32	59	C 59
Sequence 164246, A	US-09-949-016-164246	4	13.7	601	32	60	C 60
Sequence 164318, A	US-09-949-016-164318	4	13.7	601	32	61	C 61
Sequence 172085, A	US-09-949-016-172085	4	13.7	601	32	62	C 62
Sequence 185125, A	US-09-949-016-185125	4	13.7	601	32	63	C 63
Sequence 185126, A	US-09-949-016-185126	4	13.7	601	32	64	C 64
Sequence 185157, A	US-09-949-016-185157	4	13.7	601	32	65	C 65
Sequence 185158, A	US-09-949-016-185158	4	13.7	601	32	66	C 66
Sequence 187447, A	US-09-949-016-187447	4	13.7	601	32	67	C 67
Sequence 205151, A	US-09-949-016-205151	4	13.7	601	32	68	C 68
Sequence 14644, A	US-09-270-767-14644	4	13.7	756	32	69	C 69
Sequence 3645, App	US-09-270-767-3645	4	13.7	820	32	70	C 70
Sequence 18927, A	US-09-270-767-18927	4	13.7	820	32	71	C 71
Sequence 239, App	US-08-956-171E-239	4	13.7	1361	32	72	C 72
Sequence 239, App	US-08-781-986A-239	4	13.7	1361	32	73	C 73
Sequence 469, App	US-09-976-594-469	4	13.7	1503	32	74	C 74
Sequence 901, App	US-09-620-312D-901	4	13.7	1732	32	75	C 75
Sequence 231, App	US-09-976-594-231	4	13.7	1953	32	76	C 76
Sequence 5797, App	US-09-949-016-5797	4	13.7	2251	32	77	C 77
Sequence 185, App	US-09-311-021-185	4	13.7	2520	32	78	C 78
Sequence 4, Appl	US-08-100-247-4	1	13.7	2740	32	79	C 79
Sequence 4, Appl	US-08-483-146A-4	1	13.7	2740	32	80	C 80
Sequence 5, Appl	US-08-233-513A-5	1	13.7	2740	32	81	C 81
Sequence 4, Appl	US-08-484-594A-4	1	13.7	2740	32	82	C 82
Sequence 4, Appl	US-09-076-258A-4	1	13.7	2740	32	83	C 83
Sequence 4, Appl	US-08-756-031-4	2	13.7	2740	32	84	C 84
Sequence 2, Appl	US-08-897-443-2	2	13.7	3373	32	85	C 85

c 86 32 13.7 3438 4 US-09-949-016-5648 Sequence 5648, Ap
c 87 32 13.7 3438 4 US-09-949-016-5649 Sequence 5649, Ap
c 88 32 13.7 3449 4 US-09-907-794A-33 Sequence 33, Appl
c 89 32 13.7 3449 4 US-09-905-125A-33 Sequence 33, Appl
c 90 32 13.7 3449 4 US-09-902-775A-33 Sequence 33, Appl
c 91 32 13.7 3449 4 US-09-906-700-33 Sequence 33, Appl
c 92 32 13.7 3449 4 US-09-903-603A-33 Sequence 33, Appl
c 93 32 13.7 3449 4 US-09-904-920A-33 Sequence 33, Appl
c 94 32 13.7 3449 4 US-09-909-064-33 Sequence 33, Appl
c 95 32 13.7 3449 4 US-09-905-381A-33 Sequence 33, Appl
c 96 32 13.7 3449 4 US-09-906-618-33 Sequence 33, Appl
c 97 32 13.7 3496 4 US-09-949-016-344 Sequence 344, App
c 98 32 13.7 4335 4 US-09-949-016-15632 Sequence 15632, A
c 99 32 13.7 10840 4 US-09-949-016-13515 Sequence 13515, A
100 32 13.7 15148 4 US-09-949-016-15058 Sequence 15058, A

ALIGNMENTS

RESULT 1
US-09-949-016-68055/c
; Sequence 68055, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68055
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-68055
Alignment Scores:
Pred. No.: 207 Length: 601
Score: 33.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 14.16% Indels: 0
DB: 4 Gaps: 0
SEQ139-143-144-146-147-17 (1-103) x US-09-949-016-68055 (1-601)
Qy 82 PheTrp***ThrGlu 87
Db 407 TTTTGGGCGACAG 390
RESULT 2
US-09-949-016-72934/c
; Sequence 72934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72934
Alignment Scores:
Pred. No.: 207 Length: 601
Score: 33.00 Matches: 6
Percent Similarity: 28.00% Conservative: 1
Best Local Similarity: 24.00% Mismatches: 18
Query Match: 14.16% Indels: 0
DB: 4 Gaps: 0
SEQ139-143-144-146-147-17 (1-103) x US-09-949-016-72934 (1-601)
Qy 63 Asp***Leu*****Trp*****Phe 82
Db 533 GATTCAATCATCTCTTATGAAAGAGCCCTTGCAATTTCTTCCACTGATTTT 474
Qy 83 PheTrp***ThrGlu 87
Db 473 TTTTGGGCGAGGAG 459
RESULT 4
US-09-949-016-206055/c

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72934
Alignment Scores:
Pred. No.: 207 Length: 601
Score: 33.00 Matches: 6
Percent Similarity: 28.00% Conservative: 1
Best Local Similarity: 24.00% Mismatches: 18
Query Match: 14.16% Indels: 0
DB: 4 Gaps: 0
SEQ139-143-144-146-147-17 (1-103) x US-09-949-016-72934 (1-601)
Qy 63 Asp***Leu*****Trp*****Phe 82
Db 131 GATTCAATCATCTCTTATGAAAGAGCCCTTGCAATTTCTTCCACTGATTTT 72
Qy 83 PheTrp***ThrGlu 87
Db 71 TTTTGGGCGAGGAG 57
RESULT 3
US-09-949-016-72935/c
; Sequence 72935, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72935
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72935
Alignment Scores:
Pred. No.: 207 Length: 601
Score: 33.00 Matches: 6
Percent Similarity: 28.00% Conservative: 1
Best Local Similarity: 24.00% Mismatches: 18
Query Match: 14.16% Indels: 0
DB: 4 Gaps: 0
SEQ139-143-144-146-147-17 (1-103) x US-09-949-016-72935 (1-601)
Qy 63 Asp***Leu*****Trp*****Phe 82
Db 533 GATTCAATCATCTCTTATGAAAGAGCCCTTGCAATTTCTTCCACTGATTTT 474
Qy 83 PheTrp***ThrGlu 87
Db 473 TTTTGGGCGAGGAG 459
RESULT 4
US-09-949-016-206055/c

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:14 ; Search time 577.805 Seconds
(without alignments)
1055.258 Million cell updates/sec

Title: SEQ139-143-144-146-147-17

Perfect score: 233
Sequence: 1 XXXRXXPKXXXRXRXIXXXXX.....WXTXXXXXXXRXRXRXRX 103

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spo01/MYERS08974584/runat_25072005_102705_5052/app.query.fasta_1.526
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=100 -LOCAL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=MYERS08974584 @CGN 1.1 468 @runat_25072005_102705_5052 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	35.9	15.4	368	6	ABQ85527 Arabidops
2	35.9	15.4	434	10	ABX61310 Arabidops
3	35.9	15.4	639	6	ABZ14486 Arabidops
4	35	15.0	1794	12	ADJ35133 DNA encod
5	35	15.0	1794	12	ADJ35131 DNA encod

6	35	15.0	2755	10	ADE57777	Ade57777 Human gen
7	35	15.0	2755	10	ADE57781	Ade57781 Human gen
8	34.1	14.6	784	10	ADD30623	Add30623 Plant yie
9	34.1	14.6	784	12	ADI43774	Adi43774 Plant tra
c 10	34.1	14.6	1111	4	AAH29588	Aah29588 Drosophil
c 11	34	14.6	3761	8	ABZ36127	Abz36127 Human sec
c 12	33	14.2	102	10	ABX06554	Abx06554 S. pneumo
c 13	33	14.2	200	12	ADH00813	Adh00813 Kidney di
c 14	33	14.2	326	13	ADR63886	Adr63886 Cotton cd
c 15	33	14.2	376	6	ABN94533	Abn94533 Gene #103
c 16	33	14.2	422	6	ABN23546	Abn23546 Human ORF
c 17	33	14.2	488	6	ABV96225	Abv96225 Human pan
c 18	33	14.2	495	5	ABA19985	Abal9985 Human ner
c 19	33	14.2	572	13	ACN57361	Acn57361 Cotton gy
c 20	33	14.2	586	13	ADQ53130	Adq53130 Novel can
c 21	33	14.2	595	13	ADR63885	Adr63885 Cotton cd
c 22	33	14.2	649	10	ADD29745	Add29745 Mouse tum
c 23	33	14.2	652	5	ABV52638	Abv52638 Human pro
c 24	33	14.2	763	4	AAI95301	Aai95301 Human neu
c 25	33	14.2	876	2	AA330865	Aax30865 Streptoco
c 26	33	14.2	889	6	ABQ42238	Abq42238 Oligonucl
c 27	33	14.2	889	6	ABQ42239	Abq42239 Oligonucl
c 28	33	14.2	944	5	AA575089	Aas75089 DNA encod
c 29	33	14.2	1110	4	AAH90776	Aah90776 CFE 80 co
c 30	33	14.2	1113	8	ABZ42222	Abz42222 Streptoco
c 31	33	14.2	1134	4	AAH90885	Aah90885 2CFE 80 c
c 32	33	14.2	1171	13	ADS60924	Ads60924 Bacterial
c 33	33	14.2	1174	6	ABQ53835	Abq53835 Oligonucl
c 34	33	14.2	1174	6	ABQ53834	Abq53834 Oligonucl
c 35	33	14.2	1311	13	ADQ81928	Adq81928 Human BMP
c 36	33	14.2	1554	8	ABT15151	Abt15151 Pathogen
c 37	33	14.2	1575	6	ABN93157	Abn93157 Staphyloc
c 38	33	14.2	1575	13	ADS04051	Ads04051 Staphyloc
c 39	33	14.2	1603	6	ABQ60836	Abq60836 FLJ10512f
c 40	33	14.2	1624	2	AAV90773	Aav90773 Nucleotid
c 41	33	14.2	1650	4	ABL14045	Ab114045 Drosophil
c 42	33	14.2	2226	2	ABV13925	Aav13925 Human SHA
c 43	33	14.2	2269	4	ABL09273	Ab109273 Drosophil
c 44	33	14.2	2269	4	ABL15036	Ab115036 Drosophil
c 45	33	14.2	2385	4	ABL13673	Ab13673 Arabidops
c 46	33	14.2	2418	6	ABZ14230	Abz14230 Human cDN
c 47	33	14.2	2422	4	AAH14230	Aah14230 Human cDN
c 48	33	14.2	2563	2	AAV28617	Aav28617 Nucleotid
c 49	33	14.2	2563	10	ADB68454	Adb68454 Human DCA
c 50	33	14.2	2592	6	ABQ60837	Abq60837 FLJ10512f
c 51	33	14.2	2655	10	ADB68447	Adb68447 Human DCA
c 52	33	14.2	2677	13	ACN37426	Acn37426 Tumour-as
c 53	33	14.2	2677	13	ADP23044	Adp23044 PRO polyp
c 54	33	14.2	2740	3	AAC36697	Aac36697 Arabidops
c 55	33	14.2	2741	3	AAC47379	Aac47379 Arabidops
c 56	33	14.2	2761	4	ABL09154	Ab109154 Drosophil
c 57	33	14.2	2765	4	ABL28140	Ab128140 Drosophil
c 58	33	14.2	3616	2	AAV73924	Aav73924 Human SHA
c 59	33	14.2	3634	5	ABV23195	Abv23195 Human pro
c 60	33	14.2	3634	5	ABV29032	Abv29032 Human pro
c 61	33	14.2	3986	4	AAF77821	Aaf77821 Murine TS
c 62	33	14.2	4126	2	AAV90882	Aav90882 Nucleotid
c 63	33	14.2	4292	2	AAV90772	Aav90772 Nucleotid
c 64	33	14.2	4333	5	ABV28571	Abv28571 Human pro
c 65	33	14.2	4333	5	ABV22743	Abv22743 Human pro
c 66	33	14.2	4333	5	ABV27685	Abv27685 Human pro
c 67	33	14.2	4333	5	ABV21858	Abv21858 Human pro
c 68	33	14.2	4803	4	ABL09272	Ab109272 Drosophil
c 69	33	14.2	5201	12	ADP13477	Adp13477 Renal cel
c 70	33	14.2	5384	6	ABL32208	Ab132208 Human imm
c 71	33	14.2	6031	5	ABA15148	Abal5148 Human ner
c 72	33	14.2	6744	4	AAK77797	Aak77797 Human imm
c 73	33	14.2	6915	6	ABL32406	Ab132406 Human imm
c 74	33	14.2	7071	13	ADR84283	Adr84283 Aspergill
c 75	33	14.2	18627	2	AAV52246	Aav52246 Streptoco
c 76	33	14.2	44096	10	ADB68445	Adb68445 Human DCA
c 77	33	14.2	88853	11	ACN43904	Acn43904 Mouse gen
c 78	33	14.2	110000	10	AB556454_07	Continuation (8 of

79 33 14.2 110000 12 ADN46845_10 Continuation (11 o
 C 80 33 14.2 110000 12 ADN47591_09 Continuation (10 o
 C 81 33 14.2 110000 12 ADN47591_10 Continuation (11 o
 C 82 33 14.2 110000 12 ADN46123_10 Continuation (11 o
 C 83 33 14.2 110000 12 ADN47209_09 Continuation (10 o
 C 84 33 14.2 110000 12 ADN47209_10 Continuation (11 o
 C 85 33 14.2 110000 12 ADN46464_10 Continuation (11 o
 C 86 33 14.2 110000 12 ADN47960_09 Continuation (10 o
 C 87 33 14.2 110000 12 ADN47960_10 Continuation (11 o
 C 88 33 14.2 110000 13 ABD32966_02 Continuation (3 of
 C 89 33 14.2 110000 13 ABD32966_03 Continuation (4 of
 C 90 33 14.2 110000 13 ABD32806_14 Continuation (5 of
 C 91 33 14.2 110000 10 ADL13813 Adl13813 Osteoarth
 C 92 33 14.2 209484 11 ACN44126 Acn44126 Human gen
 C 93 33 14.2 210204 12 ADQ18927 Adq18927 Human sof
 C 94 33 14.2 231222 10 ADL13693 Adl13693 Osteoarth
 C 95 32.9 14.1 568 13 ACN45985 Acn45985 Cotton pr
 C 96 32.8 14.1 237 8 ABZ78105 Abz78105 Human sup
 C 97 32.8 14.1 419 10 ADI02612 Adi02612 Human cdn
 C 98 32.8 14.1 458 3 AAC16991 Aac16991 Human sec
 C 99 32.8 14.1 476 10 ADB99053 Adb99053 Lrp5 cons
 C 100 32.8 14.1 476 10 ADE82622 Ade82622 pcdna3.1

ALIGNMENTS

RESULT 1

ABQ85527 ID ABQ85527 standard; DNA; 368 BP.
 XX AC ABQ85527;
 XX AC
 XX DT 05-SEP-2002 (first entry)
 XX DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 397.
 XX DE
 XX KW Plant; Arabidopsis; transgenic; fungicide; insecticide; ds.
 XX OS Arabidopsis thaliana.
 XX OS
 XX PN US2002062014-A1.
 XX PD 23-MAY-2002.
 XX PF 26-JAN-2001; 2001US-00770791.
 XX PR 27-JAN-2000; 2000US-0178480P.
 XX PR (GORL/) GORLACH J.
 XX PA (ANY/) AN Y.
 XX PA (HAMI/) HAMILTON C M.
 XX PA (PRIC/) PRICE J L.
 XX PA (RAIN/) RAINES T M.
 XX PA (YUY/) YU Y.
 XX PA (RAME/) RAMEAKA J G.
 XX PA (PAGE/) PAGE A.
 XX PA (MATH/) MATHAW A V.
 XX PA (LEDF/) LEDFORD B L.
 XX PA (WOES/) WOESSNER J P.
 XX PA (HAAS/) HAAS W D.
 XX PA (GARC/) GARCIA C A.
 XX PA (KRICK/) KRICKER M.
 XX PA (SLAT/) SLATER T R.
 XX PA (DAVI/) DAVIS K R.
 XX PA (ALLE/) ALLEN K.
 XX PA (HOFF/) HOFFMAN N.
 XX PA (HURB/) HURBAN P.
 XX GORlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 PI Hurban P;
 XX

DR WPI; 2002-479265/51.
 XX New nucleic acid sequences of Arabidopsis thaliana and their encoded products are useful to produce transgenic plants, to screen for biologically active agents such as fungicides and insecticides and in genetic studies.
 XX Claim 1; SEQ ID NO 397; 18pp + Sequence Listing; English.
 XX The invention relates to a novel nucleic acid of Arabidopsis thaliana comprising a sequence capable of hybridising under stringency to one of the 999 sequences referred to but not defined in the specification (ABQ85131-ABQ86129). The nucleic acid sequences are useful to identify homologous or related genes, to produce compositions that modulate expression or function of the encoded protein, to map functional regions of the protein, to study associated physiological pathways, to genetically manipulate cells and plants. The encoded products are useful to screen for biologically active agents such as fungicides or insecticides and to elucidate biochemical pathways
 XX SQ Sequence 368 BP; 87 A; 100 C; 105 G; 73 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 106 Length: 368
 Score: 35.90 Matches: 7
 Percent Similarity: 26.92% Conservative: 0
 Best Local Similarity: 26.92% Mismatches: 8
 Query Match: 15.41% Indels: 11
 DB: 6 Gaps: 1

SEQ139-143-144-146-147-17 (1-103) x ABQ85527 (1-368)
 Qy 62 AspAsp***Leu*****Trr***** 81
 Db 231 GACGACTCTATCTGCACGTGG-----CGG 257
 Qy 82 PhePheTrr***ThrGlu 87
 Db 258 CTCCTGTGGACGCGGAG 275

RESULT 2
 ABX61310 ID ABX61310 standard; DNA; 434 BP.
 XX AC ABX61310;
 XX DT 26-FEB-2003 (first entry)
 XX DE Arabidopsis thaliana polynucleotide #656.
 XX KW Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;
 KW genetic modification; environmental stress; disease resistance;
 KW fungicide; insecticide; stress tolerance.
 XX OS Arabidopsis thaliana.
 XX OS
 XX PN US2002142319-A1.
 XX PD 03-OCT-2002.
 XX PF 07-AUG-2001; 2001US-00924035.
 XX PR 13-AUG-1999; 99US-0148784P.
 XX PR 11-AUG-2000; 2000US-00638258.
 XX PA (GORL/) GORLACH J.
 XX PA (ANY/) AN Y.
 XX PA (HAMI/) HAMILTON C M.
 XX PA (PRIC/) PRICE J L.
 XX PA (HARG/) HARGISS T R.
 XX PA (YUY/) YU Y.
 XX PA (RAME/) RAMEAKA J G.
 XX PA (PAGE/) PAGE A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:49 ; Search time 3386.44 Seconds
(without alignments)
1473.786 Million cell updates/sec

Title: SEQ139-143-144-146-147-17

Perfect score: 233

Sequence: 1 XXXRXXPKXXXRXRXIXXXX.....WXTXXXXXXXRXRXRXRX 103

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/MYERS08974584/runat.25072005.102706.5062/app.query.fasta_1.526
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEX=0 -LOOPEX1=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=MYERS08974584 @CN 1.1.2647 @runat.25072005.102706.5062 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	41.9	18.0	336	8	AY201359 Arabidops
C 2	39.9	17.1	185084	5	BX629350 Zebrafish
C 3	39.9	17.1	202699	2	BX640481 Danio rer
C 4	39	16.7	104128	10	AL929562 Mouse DNA

5	38	16.3	245146	2	AC130643	AC130643 Rattus no
6	38	16.3	253634	2	AC115167	AC115167 Rattus no
7	37	15.9	106208	2	AC138061	AC138061 Homo sapi
8	37	15.9	175245	10	AC129188	AC129188 Mus muscu
9	37	15.9	175542	2	AC150062	AC150062 Gallus ga
10	37	15.9	187731	10	AC140307	AC140307 Mus muscu
11	37	15.9	188771	2	AC150073	AC150073 Gallus ga
12	36.8	15.8	37841	2	AC017708	AC017708 Drosophi
13	36.8	15.8	168471	3	AC006170	AC006170 Drosophi
14	36.8	15.8	231562	3	AE003767	AE003767 Drosophi
15	36.1	15.5	4320	10	BC060679	BC060679 Mus muscu
16	36	15.5	110000	2	AC096393_5	Continuation (6 of
17	36	15.5	181561	2	AC104098	AC104098 Mus muscu
18	36	15.5	215745	10	AC117585	AC117585 Mus muscu
19	36	15.5	231612	2	AC130088	AC130088 Rattus no
20	36	15.5	245044	2	AC131471	AC131471 Rattus no
21	36	15.5	271485	2	AC137162	AC137162 Rattus no
22	36	15.5	281085	2	AC123455	AC123455 Rattus no
23	35.9	15.4	573	8	AY201051	AY201051 Arabidops
24	35.9	15.4	591	8	BT015452	BT015452 Arabidops
25	35.9	15.4	639	6	AX507596	AX507596 Sequence
26	35.9	15.4	639	8	AY074558	AY074558 Arabidops
27	35.9	15.4	702	8	AY054226	AY054226 Arabidops
28	35.9	15.4	2259	8	BT001993	BT001993 Arabidops
29	35.9	15.4	95643	8	FN15	AF069293 Arabidops
30	35.9	15.4	174269	2	AC132961	AC132961 Rattus no
31	35.9	15.4	197119	8	ATCHRIV1	AL161471 Arabidops
32	35.9	15.4	199709	2	AC113049	AC113049 Mus muscu
33	35.9	15.4	206887	2	AC134007	AC134007 Rattus no
34	35.9	15.4	259970	2	AC106985	AC106985 Rattus no
35	35.1	15.1	639	8	AF528583	AF528583 Arabidops
36	35.1	15.1	222348	10	AC111103	AC111103 Mus muscu
37	35	15.0	1611	8	AK067539	AK067539 Oryza sat
38	35	15.0	127614	8	AP005879	AP005879 Oryza sat
39	35	15.0	154912	8	AP005684	AP005684 Oryza sat
40	35	15.0	231400	2	AC109895	AC109895 Rattus no
41	34.9	15.0	705	8	AY203308	AY203308 Arabidops
42	34.9	15.0	218729	2	AC111863	AC111863 Rattus no
43	34.6	14.8	110000	2	AC151898_4	Continuation (5 of
44	34.6	14.8	110000	2	AC151898_5	Continuation (6 of
45	34.6	14.8	153108	10	AC142450	AC142450 Mus muscu
46	34.3	14.7	237624	2	AC113662	AC113662 Rattus no
47	34.3	14.7	247917	2	AC095906	AC095906 Rattus no
48	34.1	14.6	639	8	AF528584	AF528584 Arabidops
49	34.1	14.6	639	8	AF528586	AF528586 Arabidops
50	34.1	14.6	639	8	AF528588	AF528588 Arabidops
51	34.1	14.6	639	8	AF528589	AF528589 Arabidops
52	34.1	14.6	639	8	AF528590	AF528590 Arabidops
53	34.1	14.6	681	8	AY141230	AY141230 Arabidops
54	34.1	14.6	1111	6	AX094653	AX094653 Sequence
55	34.1	14.6	25749	2	AC014588	AC014588 Drosophi
56	34.1	14.6	88292	8	T8F5	AC004512 Arabidops
57	34.1	14.6	175519	3	AC010666	AC010666 Drosophi
58	34.1	14.6	250029	3	AE003599	AE003599 Drosophi
59	34	14.6	46189	8	AP006676	AP006676 Lotus cor
60	34	14.6	147109	9	AL358779	AL358779 Human DNA
61	34	14.6	149464	10	AP006583	AP006583 Gloebact
62	34	14.6	159119	1	AP006583	AC137096 Mus muscu
63	34	14.6	159552	9	AC007252	AC007252 Homo sapi
64	34	14.6	160421	2	AC025005	AC025005 Homo sapi
65	34	14.6	173062	9	AL445587	AL445587 Human DNA
66	34	14.6	175639	10	AC133515	AC133515 Mus muscu
67	34	14.6	177675	10	AC124534	AC124534 Mus muscu
68	34	14.6	181731	2	AC102525	AC102525 Mus muscu
69	34	14.6	183772	9	AC021006	AC021006 Homo sapi
70	34	14.6	206444	2	AC113538	AC113538 Mus muscu
71	34	14.6	217330	10	AC118414	AC118414 Rattus no
72	34	14.6	221697	2	AC102395	AC102395 Mus muscu
73	34	14.6	223437	2	AC110660	AC110660 Rattus no
74	34	14.6	237289	2	AC095523	AC095523 Rattus no
75	34	14.6	238352	2	AC114102	AC114102 Rattus no
76	34	14.6	238749	2	AC110965	AC110965 Rattus no
77	34	14.6	240123	2	AC098494	AC098494 Rattus no

```

c 78 33.9 14.5 28054 3 CEY40B1A
79 33.8 14.5 216813 2 AC148423
80 33.8 14.5 240857 2 AC148371
81 33.1 14.2 152748 9 AL1354777
82 33.1 14.2 155539 9 AL138721
c 83 33 14.2 102 6 AX568476
c 84 33 14.2 248 9 HG505055
c 85 33 14.2 282 6 CQ742132
c 86 33 14.2 316 3 AY427333
87 33 14.2 320 3 AY427365
c 88 33 14.2 336 6 CQ745819
89 33 14.2 376 6 AX408384
90 33 14.2 422 6 CQ449809
91 33 14.2 494 11 G31161
92 33 14.2 620 9 AF29307782
c 93 33 14.2 652 6 CQ520790
94 33 14.2 665 10 BC028521
95 33 14.2 763 6 BD019138
96 33 14.2 763 6 BD099076
c 97 33 14.2 947 9 BC042148
c 98 33 14.2 1080 9 HSDGCR6
99 33 14.2 1110 6 AX194095
100 33 14.2 1134 6 AX194317

```

ALIGNMENTS

```

RESULT 1
AY201359/c 336 bp DNA linear PLN 24-JAN-2003
LOCUS Arabidopsis thaliana sequence flanking Ds5 end of Ds-Trap Insertion
DEFINITION Arabidopsis thaliana sequence flanking Ds5 end of Ds-Trap Insertion
ACCESSION AY201359
VERSION AY201359.1 GI:27897313
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 336)
AUTHORS May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R.,
McCombie,W.R. and Martienssen,R.A.
TITLE Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 336)
AUTHORS May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R.,
McCombie,W.R. and Martienssen,R.A.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
COMMENT http://genetrap.cshl.org.
FEATURES
Location/Qualifiers
source 1..336
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GT260.Ds5.04.01.99.b.336"
/ecotype="Landsberg"
/note="transgenic line GT260"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.453 Length: 336
Score: 41.90 Matches: 8
Percent Similarity: 30.77% Conservative: 0
Best Local Similarity: 30.77% Mismatches: 7
Query Match: 17.98% Indels: 11
DB: 8 Gaps: 1

```

```

Seq139-143-144-146-147-17 (1-103) x AY201359 (1-336)

```

```

Oy 62 AspAsp***Leu*****Trr*****
Db 235 GACGACTCTCTATCTGCTACGTGG-----CGG 209
Oy 82 PhePheTrr***ThrGlu 87
Db 208 TTCTTGTGGACGACGGAG 191
RESULT 2
BX629350/c 185084 bp DNA linear VRT 19-MAR-2004
LOCUS Zebrafish DNA sequence from clone CH211-155E13 in linkage group 10,
DEFINITION complete sequence.
ACCESSION BX629350
VERSION BX629350.6 GI:45581081
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 185084)
AUTHORS Henderson,C.
TITLE Direct Submission
JOURNAL Submitted (19-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 16, 2004 this sequence version replaced gi:45379234.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep/Clone-derived
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhairong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, Washu). For further information see
http://www.sanger.ac.uk/projects/D_rerio/fishmask.shtml
CH211-155E13 is from a CHORI-211 BAC library
VECTOR: pIARBAC2.1.
Location/Qualifiers
1..185084
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-155E13"
/clone_lib="CHORI-211"

```

FEATURES

```

source

```

ORIGIN

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:18:34 ; Search time 4138.71 Seconds
(without alignments)
938.107 Million cell updates/sec

Title: SEQ139-143-144-146-147-16

Perfect score: 232

Sequence: 1 XXXXXPPXXXXXXIXXXXX.....FWTXXXXXXXXXXXXXX 102

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cn2_1/USPTO spo01/MYERS08974584/runat_25072005_102706_5075/app_query.fasta_1.526
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MYERS08974584 @CN 1 1.3556 @runat_25072005_102706_5075 -NCPU=6 -ICPU=3
-NO.MMAP -LARGUEURY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gse1:*
- 9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
C 1	39	16.8	352	5	BP427888
C 2	39	16.8	437	5	BP427888
C 3	39	16.8	576	5	BUS06243
C 4	39	16.8	599	4	BP760208
C 5	39	16.8	601	5	BI964368
C 6	39	16.8	642	5	BP766928
C 7	39	16.8	673	9	BP771659
C 8	39	16.8	707	6	CL186598
C 9	38	16.4	895	2	CD741981
C 10	38	16.4	895	2	BE617363

10	38	16.4	1047	5	BQ224762
11	37	15.9	939	9	CL398497
12	36.8	15.9	851	5	BUS09508
13	36.4	15.7	1058	8	BH725011
14	36.2	15.6	181	9	CG784907
15	36.2	15.6	197	9	CG887155
16	36.2	15.6	216	6	CB178950
17	36.2	15.6	357	9	CG784906
18	36.2	15.6	408	7	CK225321
19	36.2	15.6	563	7	CK222398
20	36.2	15.6	578	7	CN717076
21	36.2	15.6	624	7	CN699623
22	36.2	15.6	665	2	BB616800
23	36.2	15.6	741	7	CF733575
24	36.2	15.6	746	7	CK780407
25	36.2	15.6	748	6	CD803163
26	36.2	15.6	768	7	CO433173
27	36.2	15.6	859	7	CO807249
28	36.2	15.6	3232	3	AK030089
29	36.2	15.6	4003	3	AK028324
30	36.2	15.6	4024	3	AK030951
31	36.2	15.6	4306	3	AK032405
32	36.2	15.6	4942	3	AK031688
33	35.9	15.5	548	7	W43412
34	35.9	15.5	598	5	BF561807
35	35.9	15.5	651	3	CNS0A2QM
36	35.9	15.5	680	5	BX836890
37	35.9	15.5	681	1	AV822745
38	35.9	15.5	692	3	CNS0A2P9
39	35.9	15.5	693	3	CNS0YXGP
40	35.9	15.5	753	8	AZ869871
41	35.9	15.5	758	3	CNS0A2LE
42	35.9	15.5	893	7	CF652848
43	35.9	15.5	900	7	CV242818
44	35.9	15.5	928	9	CL514461
45	35.3	15.2	563	9	AX989710
46	35.1	15.1	582	1	AJ799564
47	35.1	15.1	716	1	AJ560032
48	35.1	15.1	1107	5	BQ618909
49	35.1	15.1	1107	5	BQ619092
50	35.1	15.1	1107	5	BQ619227
51	35.1	15.1	1107	5	BQ619307
52	35.1	15.1	1107	5	BQ619325
53	35	15.1	591	4	BG840829
54	35	15.1	669	8	AQ840198
55	35	15.1	761	4	BG839900
56	35	15.1	794	7	CK305937
57	34.9	15.0	449	8	AZ216917
58	34.9	15.0	554	9	TA271D10Q
59	34.2	14.7	648	8	CB461221
60	34.1	14.7	206	8	BZ660880
61	34	14.7	148	1	AA514345
62	34	14.7	229	1	AI335300
63	34	14.7	239	1	AI335291
64	34	14.7	263	7	H09046
65	34	14.7	278	5	BW171513
66	34	14.7	278	5	BW171535
67	34	14.7	294	5	BW300646
68	34	14.7	297	2	AW303078
69	34	14.7	323	7	HI8423
70	34	14.7	365	2	BF395855
71	34	14.7	381	1	AA969353
72	34	14.7	409	5	BM923188
73	34	14.7	416	2	BF920885
74	34	14.7	420	1	AI635438
75	34	14.7	424	8	AQ084051
76	34	14.7	429	2	BF409184
77	34	14.7	450	5	BX480972
78	34	14.7	450	8	A2129520
79	34	14.7	459	1	A3724649
80	34	14.7	462	1	AA861318
81	34	14.7	462	1	AI003578
82	34	14.7	493	6	CA432939

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 00:55:25 ; Search time 1176.73 Seconds
(without alignments)
560.502 Million cell updates/sec

Title: SEQ139-143-144-146-147-16
Perfect score: 232
Sequence: 1 XXXRXXPKXXXXXIXXXXXX.....FWXTXXXXXXXKXKX 102

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.1
Ygapop 10.0, Ygapext 0.1
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Total number of hits satisfying chosen parameters: 14555652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model -USER=MYERS08974584 @CN1.1.456 @runat 25072005 102708 5180/app query.fasta_1.526
-Q=/cgn2_1/USPTO.spool/USERS08974584 -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-DB=PublishedApplications_NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Block62
-TRANS=human40.cdi -LIST=100 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=BLOCK -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000
-USER=MYERS08974584 @CN1.1.456 @runat 25072005 102708 5180 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	35.9	15.5	368	9	US-09-770-791-397	Sequence 397, App
2	35.9	15.5	434	9	US-09-924-035A-656	Sequence 656, App
3	35.9	15.5	639	9	US-09-938-842A-2291	Sequence 2291, App
4	35.9	15.5	639	11	US-09-938-842A-2291	Sequence 2291, App
5	35.9	15.5	2321	20	US-10-739-930-1097	Sequence 1097, App
6	35	15.1	455	19	US-10-437-963-9511	Sequence 9511, App
7	35	15.1	600	22	US-10-972-079-62547	Sequence 62547, App
8	35	15.1	829	20	US-10-425-115-42905	Sequence 42905, App
9	34.1	14.7	784	17	US-10-225-066A-655	Sequence 655, App
10	34.1	14.7	784	17	US-10-374-780A-2237	Sequence 2237, App
11	34	14.7	404	18	US-10-424-599-40084	Sequence 40084, App
12	34	14.7	540	22	US-10-972-079-22314	Sequence 22314, App
13	34	14.7	1002	18	US-10-424-599-102834	Sequence 102834, App
14	34	14.7	1383	19	US-10-437-963-100022	Sequence 100022, App
15	33	14.2	102	21	US-10-472-928-1683	Sequence 1683, App
16	33	14.2	140	9	US-09-783-590-278	Sequence 278, App
17	33	14.2	193	19	US-10-437-963-102089	Sequence 102089, App
18	33	14.2	201	20	US-10-719-993-25917	Sequence 25917, App
19	33	14.2	282	18	US-10-424-599-57538	Sequence 57538, App
20	33	14.2	326	19	US-10-767-795-4667	Sequence 4667, App
21	33	14.2	330	9	US-09-783-590-11246	Sequence 11246, App
22	33	14.2	376	9	US-09-880-107-1031	Sequence 1031, App
23	33	14.2	460	20	US-10-425-115-80762	Sequence 80762, App
24	33	14.2	488	14	US-10-060-036-1633	Sequence 1633, App
25	33	14.2	521	20	US-10-425-115-36983	Sequence 36983, App
26	33	14.2	529	19	US-10-767-701-5515	Sequence 5515, App
27	33	14.2	539	18	US-10-424-599-29238	Sequence 29238, App
28	33	14.2	572	19	US-10-021-323-12142	Sequence 12142, App
29	33	14.2	593	13	US-10-027-632-271755	Sequence 271755, App
30	33	14.2	593	17	US-10-027-632-271755	Sequence 271755, App
31	33	14.2	595	13	US-10-767-795-4666	Sequence 4666, App
32	33	14.2	616	13	US-10-027-632-265805	Sequence 265805, App
33	33	14.2	616	13	US-10-027-632-265806	Sequence 265806, App
34	33	14.2	616	17	US-10-027-632-265805	Sequence 265805, App
35	33	14.2	616	17	US-10-027-632-265806	Sequence 265806, App
36	33	14.2	619	13	US-10-027-632-196891	Sequence 196891, App
37	33	14.2	619	13	US-10-027-632-196892	Sequence 196892, App
38	33	14.2	619	13	US-10-027-632-196893	Sequence 196893, App
39	33	14.2	619	17	US-10-027-632-196891	Sequence 196891, App
40	33	14.2	619	17	US-10-027-632-196892	Sequence 196892, App
41	33	14.2	619	17	US-10-027-632-196893	Sequence 196893, App
42	33	14.2	649	20	US-10-335-053-194	Sequence 194, App
43	33	14.2	652	20	US-10-357-930-52657	Sequence 52657, App
44	33	14.2	664	18	US-10-425-114-16192	Sequence 16192, App
45	33	14.2	667	13	US-10-027-632-287415	Sequence 287415, App
46	33	14.2	667	13	US-10-027-632-287416	Sequence 287416, App
47	33	14.2	667	17	US-10-027-632-287415	Sequence 287415, App
48	33	14.2	667	17	US-10-027-632-287416	Sequence 287416, App
49	33	14.2	721	13	US-10-027-632-260266	Sequence 260266, App
50	33	14.2	721	17	US-10-027-632-260266	Sequence 260266, App
51	33	14.2	900	20	US-10-425-115-160420	Sequence 160420, App
52	33	14.2	987	18	US-10-424-599-90405	Sequence 90405, App
53	33	14.2	1113	19	US-10-474-776-76	Sequence 76, App
54	33	14.2	1162	21	US-10-956-157-1455	Sequence 1455, App
55	33	14.2	1162	17	US-10-956-157-6690	Sequence 6690, App
56	33	14.2	1171	17	US-10-369-493-36598	Sequence 36598, App
57	33	14.2	1171	20	US-10-363-345A-40425	Sequence 40425, App
58	33	14.2	1174	20	US-10-363-345A-40426	Sequence 40426, App
59	33	14.2	1174	21	US-10-363-483A-40426	Sequence 40426, App
60	33	14.2	1174	21	US-10-363-483A-40426	Sequence 40426, App
61	33	14.2	1201	20	US-10-425-115-89303	Sequence 89303, App
62	33	14.2	1312	18	US-10-425-114-25327	Sequence 25327, App
63	33	14.2	1404	20	US-10-425-115-49114	Sequence 49114, App
64	33	14.2	1554	21	US-10-470-048B-537	Sequence 537, App

```
65 33 14.2 1575 22 US-10-724-972A-3346
C 66 33 14.2 1631 13 US-10-027-632-258931
C 67 33 14.2 1631 13 US-10-027-632-258932
C 68 33 14.2 1631 13 US-10-027-632-258933
C 69 33 14.2 1631 13 US-10-027-632-258934
C 70 33 14.2 1631 17 US-10-027-632-258931
C 71 33 14.2 1631 17 US-10-027-632-258932
C 72 33 14.2 1631 17 US-10-027-632-258933
C 73 33 14.2 1631 17 US-10-027-632-258934
C 74 33 14.2 1988 13 US-10-027-632-97163
C 75 33 14.2 1988 13 US-10-027-632-97164
C 76 33 14.2 1988 13 US-10-027-632-97165
C 77 33 14.2 1988 17 US-10-027-632-97163
C 78 33 14.2 1988 17 US-10-027-632-97164
C 79 33 14.2 1988 17 US-10-027-632-97165
C 80 33 14.2 2418 9 US-09-938-842A-1478
C 81 33 14.2 2418 11 US-09-938-842A-1478
C 82 33 14.2 2563 9 US-09-782-051-1
C 83 33 14.2 2655 21 US-10-876-285-3
C 84 33 14.2 2760 20 US-10-739-930-157
C 85 33 14.2 3634 20 US-10-357-930-23184
C 86 33 14.2 3634 20 US-10-357-930-29050
C 87 33 14.2 3986 13 US-10-092-925-1
C 88 33 14.2 4105 17 US-10-062-674-1991
C 89 33 14.2 4333 20 US-10-357-930-21849
C 90 33 14.2 4333 20 US-10-357-930-22736
C 91 33 14.2 4333 20 US-10-357-930-27699
C 92 33 14.2 4333 20 US-10-357-930-28584
C 93 33 14.2 5201 19 US-10-717-597-213
C 94 33 14.2 5384 15 US-10-311-455-181
C 95 33 14.2 6515 15 US-10-311-455-379
C 96 33 14.2 18627 8 US-08-961-527-113
C 97 33 14.2 18627 17 US-10-158-844-113
C 98 33 14.2 44096 21 US-10-876-285-1
C 99 33 14.2 88853 13 US-10-087-192-85
C 100 33 14.2 127602 22 US-10-737-082-65

ALIGNMENTS

RESULT 1
US-09-770-791-397
; Sequence 397, Application US/0970791
; Patent No. US20020062014A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Kricker, Ted
; APPLICANT: Slader, David R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2029 (PARA-018PRV)
; CURRENT APPLICATION NUMBER: US/09/770,791
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,480
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0

Sequence 3346, Ap
Sequence 258931,
Sequence 258932,
Sequence 258933,
Sequence 258934,
Sequence 258931,
Sequence 258932,
Sequence 258933,
Sequence 258934,
Sequence 97163, A
Sequence 97164, A
Sequence 97165, A
Sequence 97163, A
Sequence 97164, A
Sequence 97165, A
Sequence 1478, Ap
Sequence 1478, Ap
Sequence 1, Appli
Sequence 1, Appli
Sequence 157, App
Sequence 23184, A
Sequence 29050, A
Sequence 1, Appli
Sequence 1991, Ap
Sequence 21849, A
Sequence 2736, A
Sequence 27699, A
Sequence 28584, A
Sequence 213, App
Sequence 181, App
Sequence 379, App
Sequence 113, App
Sequence 113, App
Sequence 1, Appli
Sequence 85, Appli
Sequence 65, Appli

; SEQ ID NO 397
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(368)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-791-397

Alignment Scores:
Pred. No.: 183
Score: 35.90
Percent Similarity: 26.92%
Best Local Similarity: 26.92%
Query Match: 15.47%
DB: 9
Length: 368
Matches: 7
Conservative: 0
Mismatch: 8
Indels: 11
Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-09-770-791-397 (1-368)

Qy 62 AspAsp***Leu*****Trp***** 81
|||
Db 231 GACGACTCTCTATCTGCCACGTGG-----CGG 257
|||

Qy 82 PhePheTrp***ThrGlu 87
|||
Db 258 CTCTTGTGGACGACGGAG 275
|||

RESULT 2
US-09-924-035A-656
; Sequence 656, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grisch, Jrn
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-13
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 656
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(434)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-656

Alignment Scores:
Pred. No.: 216
Score: 35.90
Percent Similarity: 26.92%
Best Local Similarity: 26.92%
Query Match: 15.47%
DB: 9
Length: 434
Matches: 7
Conservative: 0
Mismatch: 8
Indels: 11
Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-09-924-035A-656 (1-434)

Qy 62 AspAsp***Leu*****Trp***** 81
|||
Db 293 GACGACTCTCTATCTGCCACGTGG-----CGG 319
|||

Qy 82 PhePheTrp***ThrGlu 87
|||
Db 320 CTCTTGTGGACGACGGAG 337
|||

RESULT 3
US-09-938-842A-2291
```



```
; Sequence 2291, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2291
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2291

Alignment Scores:
Pred. No.: 322 Length: 639
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservativeness: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 9 Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-09-938-842A-2291 (1-639)

Qy 62 AspAsp***Leu*****Trp***** 81
Db 241 GACGACTCTCTATCTGCCACGTGG-----CGG 267

Qy 82 PhePheTrp***ThrGlu 87
Db 268 CTCCTGTGGACGACGGAG 285

RESULT 4
US-10-739-930-1097
; Sequence 1097, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 1097
; LENGTH: 2321
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER6232_1
US-10-739-930-1097

Alignment Scores:
Pred. No.: 1,21e+03 Length: 2321
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservativeness: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 20 Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-10-739-930-1097 (1-2321)

Qy 62 AspAsp***Leu*****Trp***** 81
Db 280 GACGACTCTCTATCTGCCACGTGG-----CGG 306

Qy 82 PhePheTrp***ThrGlu 87
Db 307 CTCCTGTGGACGACGGAG 324

RESULT 6
US-10-437-963-9511/c
; Sequence 9511, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 9511
```

```
; Sequence 2291, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2291
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2291

Alignment Scores:
Pred. No.: 322 Length: 639
Score: 35.90 Matches: 7
Percent Similarity: 26.92% Conservativeness: 0
Best Local Similarity: 26.92% Mismatches: 8
Query Match: 15.47% Indels: 11
DB: 9 Gaps: 1

SEQ139-143-144-146-147-16 (1-102) x US-09-938-842A-2291 (1-639)

Qy 62 AspAsp***Leu*****Trp***** 81
Db 241 GACGACTCTCTATCTGCCACGTGG-----CGG 267

Qy 82 PhePheTrp***ThrGlu 87
Db 268 CTCCTGTGGACGACGGAG 285

RESULT 4
US-09-938-842A-2291
; Sequence 2291, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2291
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2291

Alignment Scores:
Pred. No.: 322 Length: 639
Score: 35.90 Matches: 7
```

```
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_15921C.1
;   US-10-437-963-9511
Alignment Scores:
Pred. No.:      362      Length:      455
Score:          35.00    Matches:      6
Percent Similarity: 31.58% Conservative: 0
Best Local Similarity: 31.58% Mismatches: 13
Query Match:      15.09% Indels:      0
DB:              19      Gaps:      0

SEQ139-143-144-146-147-16 (1-102) x US-10-437-963-9511 (1-455)
Qy 69 Trp*****PhePheTrp***ThrGlu 87
Db 118 TGGATGGATAGTGGTCCCTTATCTGTCCTCTTTTCTTGGAAACACAG 62

RESULT 7
US-10-972-079-62547
; Sequence 62547, Application US/10972079
; Publication No. US2005015317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEER
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; PRIOR FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62547
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894324709_1
US-10-972-079-62547
Alignment Scores:
Pred. No.:      481      Length:      600
Score:          35.00    Matches:      6
Percent Similarity: 27.27% Conservative: 0
Best Local Similarity: 27.27% Mismatches: 16
Query Match:      15.09% Indels:      0
DB:              22      Gaps:      0

SEQ139-143-144-146-147-16 (1-102) x US-10-972-079-62547 (1-600)
Qy 65 Leu*****Trp*****PhePheTrp 84
Db 65 CTGCAGTCTCTGGTGTCTACCTTCAGGATCTCCAGCATTCATCTGCTCTCTCTGG 124

Qy 85 ***Thr 86
Db 125 GAAACG 130

RESULT 8
US-10-425-115-42905/c
; Sequence 42905, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 42905
; LENGTH: 829
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_139137C.1
;   US-10-425-115-42905
Alignment Scores:
Pred. No.:      670      Length:      829
Score:          35.00    Matches:      6
Percent Similarity: 28.57% Conservative: 0
Best Local Similarity: 28.57% Mismatches: 15
Query Match:      15.09% Indels:      0
DB:              20      Gaps:      0

SEQ139-143-144-146-147-16 (1-102) x US-10-425-115-42905 (1-829)
Qy 82 PhePheTrp***ThrGlu*****Arg***** 101
Db 253 TCTTTTGGACGACGCCCTCACTGGCGCCGGGCGGATGACTCGGAGAGAGACCTTGGGTC 194

Qy 102 Trp 102
Db 193 TGG 191

RESULT 9
US-10-225-066A-655
; Sequence 655, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 655
; LENGTH: 784
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-655
```


C 86	32	13.8	2740	4	US-08-756-031-4	Sequence 4, Appli
C 87	32	13.8	3373	2	US-08-897-443-2	Sequence 2, Appli
C 88	32	13.8	3438	4	US-09-949-016-5648	Sequence 5648, Ap
C 89	32	13.8	3438	4	US-09-949-016-5649	Sequence 5649, Ap
C 90	32	13.8	3449	4	US-09-907-794A-33	Sequence 33, Appl
C 91	32	13.8	3449	4	US-09-905-125A-33	Sequence 33, Appl
C 92	32	13.8	3449	4	US-09-902-775A-33	Sequence 33, Appl
C 93	32	13.8	3449	4	US-09-906-700-33	Sequence 33, Appl
C 94	32	13.8	3449	4	US-09-903-603A-33	Sequence 33, Appl
C 95	32	13.8	3449	4	US-09-904-920A-33	Sequence 33, Appl
C 96	32	13.8	3449	4	US-09-909-064-33	Sequence 33, Appl
C 97	32	13.8	3449	4	US-09-903-381A-33	Sequence 33, Appl
C 98	32	13.8	3449	4	US-09-906-618-33	Sequence 33, Appl
C 99	32	13.8	3496	4	US-09-949-016-344	Sequence 344, App
C 100	32	13.8	4335	4	US-09-949-016-15632	Sequence 15632, A

ALIGNMENTS

```

RESULT 1
US-09-949-016-25892
; Sequence 25892, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J, Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25892
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-25892

```

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61322
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-61322

Alignment Scores:
Pred. No.:      24.8      Length:      601
Score:          37.00     Matches:      6
Percent Similarity: 28.57% Conservative: 0
Best Local Similarity: 28.57% Mismatches: 15
Query Match:      15.95% Indels:      0
DB:               4       Gaps:       0

SEQ139-143-144-146-147-16 (1-102) x US-09-949-016-61322 (1-601)

Qy      82 PhePheTrp***ThrGlu*****Arg***** 101
Db      313 TTTTITTTTGAAGACAGAATCTCGCTCTGTTGCCCAGGCTGGAGTGCAGCGCAGCATCT 372

Qy      102 Trp 102
Db      373 TGG 375

RESULT 3
US-09-949-016-12144
; Sequence 12144, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:14 ; Search time 572.195 Seconds
(without alignments)
1055.258 Million cell updates/sec

Title: SEQ139-143-144-146-147-16

Perfect score: 232

Sequence: 1 XXXRXXPKXXXXXRXIXXXXX.....FWXTXXXXXXXRXRXW 102

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.1
Ygapop 10.0 , Ygapext 0.1
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlh
-Q/cgn2_1/USPTO.spool/WERS08974584/runat_25072005_102705_5052/app_query.fasta_1.526
-DB=N Geneseq_16Dec04 -QWMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blowum62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=WERS08974584 @CGN 1 1 468 @runat_25072005_102705_5052 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

N Geneseq_16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	37	15.9	25402	4	Aak68523 Human imm
2	35.9	15.5	368	6	Abq85527 Arabidops
3	35.9	15.5	434	10	Abx61310 Arabidops
4	35.9	15.5	639	6	Abz14486 Arabidops
5	35	15.1	2755	10	Ade57777 Human gen

6	35	15.1	2755	10	ADES7781	Ade57781 Human gen
7	34.1	14.7	784	10	ADD30623	Ad30623 Plant yie
8	34.1	14.7	784	12	ADI43774	Adi43774 Plant tra
9	34.1	14.7	1111	4	AAH29588	Aah29588 Drosophil
10	34	14.7	388	4	AAI91255	Aai91255 Human pol
11	34	14.7	570	10	ADF90573	Adf90573 Human hep
12	34	14.7	784	4	AAK51974	Aak51974 Human pol
13	34	14.7	827	4	AAK52958	Aak52958 Human pol
14	34	14.7	1469	3	AAC68082	Aac68082 Human sec
15	34	14.7	3761	8	ABZ36127	Abz36127 Human sec
16	33.1	14.3	2269	4	ABL09273	Abi09273 Drosophil
17	33.1	14.3	2761	4	ABL09154	Abi09154 Drosophil
18	33.1	14.3	4803	4	ABL09272	Abi09272 Drosophil
19	33	14.2	102	10	ABX06554	Abx06554 S. pneumo
20	33	14.2	200	12	ADH00813	Adh00813 Kidney di
21	33	14.2	326	13	ADR63886	Adr63886 Cotton cd
22	33	14.2	376	6	ABN94533	Abn94533 Gene #103
23	33	14.2	422	6	ABN23546	Abn23546 Human ORF
24	33	14.2	488	6	ABV96225	Abv96225 Human pan
25	33	14.2	495	5	ABA19985	Abal9985 Human ner
26	33	14.2	572	13	ACN57361	Acn57361 Cotton gy
27	33	14.2	586	13	ADQ53130	Adq53130 Novel can
28	33	14.2	595	13	ADR63885	Adr63885 Cotton cd
29	33	14.2	649	10	ADD29745	Add29745 Mouse tum
30	33	14.2	652	5	ABV52638	Abv52638 Human pro
31	33	14.2	763	4	AAI95301	Aai95301 Human neu
32	33	14.2	876	2	AAK30865	Aak30865 Streptoco
33	33	14.2	1110	4	AAH90776	Aah90776 CFE 80 co
34	33	14.2	1113	8	ABZ42222	Abz42222 Streptoco
35	33	14.2	1134	4	AAH90885	Aah90885 2CFE 80 c
36	33	14.2	1171	13	ADS60924	Adm60924 Bacterial
37	33	14.2	1174	6	ABQ53835	Abq53835 Oligonuc1
38	33	14.2	1174	6	ABQ53834	Abq53834 Oligonuc1
39	33	14.2	1311	13	ADQ81928	Adq81928 Human BMP
40	33	14.2	1554	8	ABT15151	Abt15151 Pathogen
41	33	14.2	1575	6	ABN93157	Abn93157 Staphyloc
42	33	14.2	1575	13	ADS04051	Adm04051 Staphyloc
43	33	14.2	1603	6	ABQ60836	Abq60836 FLJ10512f
44	33	14.2	1624	2	AAV90773	Aav90773 Nucleotid
45	33	14.2	1650	4	ABL14045	Abi14045 Drosophil
46	33	14.2	2226	2	AAV73925	Aav73925 Human SAH
47	33	14.2	2226	10	ADF76062	Adf76062 Human DNA
48	33	14.2	2385	4	ABL15036	Abi15036 Drosophil
49	33	14.2	2418	6	ABZ13673	Abz13673 Arabidops
50	33	14.2	2422	4	AAH14230	Aah14230 Human CDN
51	33	14.2	2563	2	AAV28617	Aav28617 Nucleotid
52	33	14.2	2563	10	ADB88454	Adb88454 Human DCA
53	33	14.2	2592	6	ABQ60837	Abq60837 FLJ10512f
54	33	14.2	2655	10	ADB68447	Adb68447 Human DCA
55	33	14.2	2677	13	ACN37426	Acn37426 Tumour-as
56	33	14.2	2677	13	ADP23044	Adp23044 PRO polyp
57	33	14.2	2740	3	AAK36697	Aac36697 Arabidops
58	33	14.2	2741	3	AAK47379	Aac47379 Arabidops
59	33	14.2	2765	4	ABL28140	Abi28140 Drosophil
60	33	14.2	3616	2	AAV73924	Aav73924 Human SAH
61	33	14.2	3634	5	ABV23195	Abv23195 Human pro
62	33	14.2	3634	5	ABV29032	Abv29032 Human pro
63	33	14.2	3986	4	AAV77821	Aav77821 Murine TS
64	33	14.2	4126	4	AAV90882	Aav90882 Nucleotid
65	33	14.2	4292	2	AAV90772	Aav90772 Nucleotid
66	33	14.2	4333	5	ABV28571	Abv28571 Human pro
67	33	14.2	4333	5	ABV22743	Abv22743 Human pro
68	33	14.2	4333	5	ABV27685	Abv27685 Human pro
69	33	14.2	4333	5	ABV21858	Abv21858 Human pro
70	33	14.2	5201	12	ADP13477	Adp13477 Renal cel
71	33	14.2	5384	6	ABL32208	Abi32208 Human imm
72	33	14.2	6031	5	ABA15148	Abal5148 Human ner
73	33	14.2	6744	4	AAK77797	Aak77797 Human imm
74	33	14.2	6915	6	ABL32406	Abi32406 Human imm
75	33	14.2	7071	13	ADR42483	Adr42483 Aspergill
76	33	14.2	18627	2	AAV52246	Aav52246 Streptoco
77	33	14.2	44096	10	ADB68445	Adb68445 Human DCA
78	33	14.2	88853	11	ACN43904	Acn43904 Mouse gen

```
79 33 14.2 110000 10 ABS56454 07 Continuation (8 of
80 33 14.2 110000 12 ADNA46845_10 Continuation (11 o
81 33 14.2 110000 12 ADNA47591_09 Continuation (11 o
82 33 14.2 110000 12 ADNA47591_10 Continuation (11 o
83 33 14.2 110000 12 ADNA46123_10 Continuation (11 o
84 33 14.2 110000 12 ADNA47209_09 Continuation (10 o
85 33 14.2 110000 12 ADNA47209_10 Continuation (11 o
86 33 14.2 110000 12 ADNA46464_10 Continuation (11 o
87 33 14.2 110000 12 ADNA47960_09 Continuation (11 o
88 33 14.2 110000 12 ADNA47960_10 Continuation (11 o
89 33 14.2 110000 13 ABD32966_02 Continuation (3 of
90 33 14.2 110000 13 ABD32966_03 Continuation (4 of
91 33 14.2 110000 13 ABD32806_4 Continuation (5 of
92 33 14.2 179651 10 ADL13813 Adl13813 Osteoearth
93 33 14.2 209484 11 ACN44126 Acn44126 Human gen
94 33 14.2 210204 12 ADQ18927 Adq18927 Human sof
95 33 14.2 231222 10 ADL13693 Adl13693 Osteoearth
96 32.9 14.2 568 13 ACN45985 Acn45985 Cotton pr
97 32.8 14.1 237 8 ABZ78105 Abz78105 Human Sup
98 32.8 14.1 419 10 ADI02612 Adi02612 Human cDN
99 32.8 14.1 458 3 AAc16991 Aac16991 Human sec
100 32.8 14.1 476 10 ADB99053 Adb99053 LRP5 cons
```

ALIGNMENTS

```
RESULT 1
AAK68523
ID AAK68523 standard; DNA; 25402 BP.
XX
AC AAK68523;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23335.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytosstatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
FN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
FF 17-JAN-2001; 2001WO-US0001354.
XX
31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
```

```
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2005, 22:14:49 ; Search time 3353.56 Seconds
(without alignments)
1473.786 Million cell updates/sec

Title: SEQ139-143-144-146-147-16

Perfect score: 232

Sequence: 1 XXXXXPKKXXXXXIXXXX.....FWXTXXXXXXXRXRXW 102

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.1
Xgapop 10.0			
Xgapext 0.1			
Fgapop 6.0			
Delopt 6.0			

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/MyERS08974584/runat_25072005_102706_5062/app_query.fasta_1.526
-DB=GenEmbl -QWTF=faescap -SUFFIX=rgc -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCLALIGN=200 -THR SCORES=ptc -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=MYERS08974584 @CGN 1.1.2647 -runat_25072005_102706_5062 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBBLACK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.1 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.hcg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	41.9	18.1	336	8	AY201359 Arabidops
C 2	39.9	17.2	185084	5	BX629350 Zebrafish
C 3	39.9	17.2	202699	2	BX640481 Danio rer
C 4	39	16.8	104128	10	AL929562 Mouse DNA

AC130643	Rattus no	16.4	245146	2	AC130643
AC115167	Rattus no	16.4	253634	2	AC115167
AC138061	Homo sapi	15.9	106208	2	AC138061
AL139131	Homo sapi	15.9	145882	2	AL139131
AC019260	Homo sapi	15.9	153557	2	AC019260
AC129188	Mus muscu	15.9	175245	10	AC129188
AC150062	Gallus ga	15.9	175542	2	AC150062
AL390718	Homo sapi	15.9	176123	9	AL390718
AC140307	Mus muscu	15.9	187731	10	AC140307
AC150073	Gallus ga	15.9	188771	2	AC150073
AC134732	Rattus no	15.9	264975	2	AC134732
AC017708	Drosophil	15.9	37841	2	AC017708
AC006170	Drosophil	15.9	168471	3	AC006170
AE003767	Drosophil	15.9	231562	3	AE003767
BC060679	Mus muscu	15.9	4320	10	BC060679
Continuation (6 of		15.5	110000	2	AC096393
AC104098	Mus muscu	15.5	181561	2	AC104098
AC117585	Mus muscu	15.5	215745	10	AC117585
AC130088	Rattus no	15.5	231612	2	AC130088
AC131471	Rattus no	15.5	245044	2	AC131471
AC137162	Rattus no	15.5	271485	2	AC137162
AC123455	Rattus no	15.5	281085	2	AC123455
AY201051	Arabidops	15.5	573	8	AY201051
BT015452	Arabidops	15.5	591	8	BT015452
AX507596	Sequence	15.5	639	6	AX507596
AY074558	Arabidops	15.5	639	8	AY074558
AY054226	Arabidops	15.5	702	8	AY054226
BT001993	Arabidops	15.5	2259	8	BT001993
AF069299	Arabidops	15.5	95643	8	AF069299
AC132961	Rattus no	15.5	174269	2	AC132961
AL161471	Arabidops	15.5	197119	8	AL161471
AC113049	Mus muscu	15.5	199709	2	AC113049
AC134007	Rattus no	15.5	206887	2	AC134007
AC106985	Rattus no	15.5	259970	2	AC106985
AF528583	Arabidops	15.5	639	8	AF528583
AC111103	Oryza sat	15.1	222348	10	AC111103
AK067539	Oryza sat	15.1	1611	8	AK067539
AL365230	Human DNA	15.1	19257	8	AL365230
AP005879	Oryza sat	15.1	127614	8	AP005879
AP005884	Oryza sat	15.1	154912	8	AP005884
AC016037	Homo sapi	15.1	161222	2	AC016037
AC109895	Rattus no	15.1	231400	2	AC109895
AY203308	Arabidops	15.0	705	8	AY203308
AC111863	Rattus no	15.0	218729	2	AC111863
Continuation (5 of		14.9	110000	2	AC151898
Continuation (6 of		14.9	110000	2	AC151898
AC142450	Mus muscu	14.9	153108	10	AC142450
AC113662	Rattus no	14.8	237624	2	AC113662
AC095906	Rattus no	14.8	247917	2	AC095906
AF528584	Arabidops	14.7	639	8	AF528584
AF528586	Arabidops	14.7	639	8	AF528586
AF528588	Arabidops	14.7	639	8	AF528588
AF528589	Arabidops	14.7	639	8	AF528589
AF528590	Arabidops	14.7	639	8	AF528590
AV141230	Arabidops	14.7	681	8	AV141230
AX094653	Sequence	14.7	1111	6	AX094653
AC014588	Drosophil	14.7	25749	2	AC014588
AC004512	Arabidops	14.7	88292	2	AC004512
AP006583	Gloebact	14.7	159119	1	AP006583
AC010666	Drosophil	14.7	175519	3	AC010666
AC102395	Mus muscu	14.7	221697	2	AC102395
AC114102	Rattus no	14.7	238352	2	AC114102
AE003599	Drosophil	14.7	250029	3	AE003599
BV201245	squm20613	14.7	201	11	BV201245
AY023258	Oryza sat	14.7	224	8	AY023258
AY692737	Saccharom	14.7	1239	8	AY692737
BD252051	47 secret	14.7	1469	6	BD252051
BC063045	Homo sapi	14.7	1700	9	BC063045
AK073936	Oryza sat	14.7	2065	8	AK073936
Z73162	S.cerevisia	14.7	2112	8	Z73162
AK112076	Oryza sat	14.7	2121	8	AK112076
AK111818	Oryza sat	14.7	2353	8	AK111818
BC050388	Homo sapi	14.7	3256	9	BC050388

```

78 34 14.7 10849 1 AB011567
79 34 14.7 37396 8 SCCH13LST
80 34 14.7 46189 8 AF006676
81 34 14.7 50550 7 AF271693
82 34 14.7 100899 9 AC026436
83 34 14.7 108098 9 AC008411
84 34 14.7 110000 1 AC008411
85 34 14.7 114721 8 AP004094
86 34 14.7 132906 8 CNS08C9H
87 34 14.7 139277 8 AF005115
88 34 14.7 145890 8 CNS08C7O
89 34 14.7 147369 2 AC147689
90 34 14.7 147512 9 AC074020
91 34 14.7 159952 9 AC007252
92 34 14.7 175639 10 AC133515
93 34 14.7 177675 10 AC124534
94 34 14.7 179821 2 AC147950
95 34 14.7 183772 9 AC021006
96 34 14.7 192191 2 AC150426
97 34 14.7 203638 2 AC135953
98 34 14.7 217330 10 AC148414
99 34 14.7 223437 2 AC110660
100 34 14.7 237289 2 AC095523

```

ALIGNMENTS

```

RESULT 1
AY201359/c
LOCUS
DEFINITION Arabidopsis thaliana sequence flanking Ds5 end of Ds-Trap insertion
from line GT260.
ACCESSION AY201359
VERSION AY201359.1 GI:27897313
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 336)
May,B.P., Sinorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R.,
McCombie,W.R. and Martienssen,R.A.
Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
Unpublished
2 (bases 1 to 336)
May,B.P., Sinorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R.,
McCombie,W.R. and Martienssen,R.A.
Direct Submission
Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
http://genetrap.cshl.org.
Location/Qualifiers
1. 336
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GT260.DS5.04.01.99.b.336"
/ecotype="Landsberg"
/note="transgenic line GT260"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 0.48 Length: 336
Score: 41.90 Matches: 8
Percent Similarity: 30.77% Conservative: 0
Best Local Similarity: 30.77% Mismatches: 7
Query Match: 18.06% Indels: 11
DB: 8 Gaps: 1
SQ139-143-144-146-147-16 (1-102) x AY201359 (1-336)

```

```

Qy 62 AspAap**Leu*****Trr***** 185084 bp DNA linear VRT 19-MAR-2004
||| ||| ||| |||
235 GACGACTCTCTAATCTGCTAGTGG-----CGG 209

Qy 82 PhePheTrr**ThrGlu 87
||| ||| ||| |||
208 TTCTTGTGGACGACGGAG 191

RESULT 2
BX629350/c
LOCUS
DEFINITION Zebrafish DNA sequence from clone CH211-155E13 in linkage group 10,
complete sequence.
ACCESSION BX629350
KEYWORDS BX629350.6 GI:45581081
VERSION HTG.
SOURCE BX629350
ORGANISM Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 185084)
Henderson,C.
Direct Submission
Submitted (19-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 18, 2004 this sequence version replaced gi:45379234.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/c_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-155E13 is from a CHORI-211 BAC library

FEATURES

```

Location/Qualifiers
1. 185084
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-155E13"
/clone_lib="CHORI-211"

```

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 04:04:20 ; Search time 85 Seconds
(without alignments)
6819.690 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVNSLLRSYRE.....TALEAANPALPDSFKTILD 1132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5952	99.8	1132	1	TERT HUMAN
2	5574.5	93.5	1069	2	Q8NG46
3	4080.5	68.4	1123	1	TERT CANFA
4	4046	67.9	807	2	Q8N6C3
5	3667	66.5	795	2	Q8NG38
6	3619	60.7	1128	2	Q9QXZ4
7	3520.5	59.0	1125	2	Q673L6
8	3505.5	58.8	1119	2	Q673L5
9	3496	58.6	1122	1	TERT MOUSE
10	2825	47.4	524	2	Q9UBR6
11	2581	43.3	1346	2	Q6RD80
12	2309.5	38.7	1191	2	Q9DE32
13	2207.5	37.0	523	2	Q94807
14	2007	33.7	575	2	Q9JK99
15	1408	23.6	615	2	Q673L3
16	1373	23.0	514	2	Q9R266
17	717.5	12.0	1123	2	Q9SE99
18	717.5	12.0	1123	2	Q9SPU7
19	683	11.5	1259	2	Q9AUI3
20	649.5	10.9	1261	2	Q8LKW0
21	588	9.9	988	1	TERT SCHPO
22	469	7.9	1032	2	Q7Z1L1
23	468.5	7.9	939	2	Q7SD71
24	465	7.8	1032	2	Q8MUB3
25	450	7.5	85	2	Q9UN56
26	449	7.5	116	2	Q80SU5
27	447	7.5	940	2	Q6C9D0
28	440.5	7.4	1031	1	TERT_EUPAE
29	432	7.2	1032	2	Q7Z1L0
30	404	6.8	823	2	Q8SQQ0
31	397	6.7	104	2	Q9JLM1

ALIGNMENTS

32	394	6.6	73	2	Q9UNR4	Q9unr4 homo sapien
33	391.5	6.6	861	2	Q6CSS0	Q6cso0 kluyveromy
34	388.5	6.5	1135	2	Q818Z7	Q818z7 sterkiella
35	388	6.5	1132	1	TERT_OXYTR	O76332 oxytricha t
36	385	6.5	1108	2	Q818Z6	Q818z6 sterkiella
37	366	6.1	867	2	Q9P8T3	Q9p8t3 candida alb
38	362	6.1	867	2	Q9P8T2	Q9p8t2 candida alb
39	359	6.0	895	2	Q9GRCS	Q9grcs paramescium
40	357	6.0	1117	1	TERT_TETTH	O77448 tetrahymena
41	355.5	6.0	884	1	TERT_YEAST	Q06163 saccharomyce
42	351.5	5.9	894	2	Q6BUF6	Q6buf6 debaryomyce
43	339	5.7	79	2	Q7YR69	Q7yr69 felis silve
44	338	5.7	79	2	Q76K45	Q76k45 canis famil
45	337	5.7	896	2	Q8MUQ8	Q8muq8 paramescium
46	318	5.3	853	2	Q752S9	Q752s9 ashbya goss
47	310	5.2	851	2	Q6FXX4	Q6fxx4 candida gla
48	244.5	4.1	261	2	Q67EQ2	Q67eq2 hordeum vul
49	195	3.3	260	2	Q6W8T6	Q6w8t6 euplotes ra
50	194	3.3	52	2	Q9R0B3	Q9r0b3 mus musculu
51	191	3.2	233	2	Q6SXU4	Q6sxu4 euplotes mi
52	191	3.2	233	2	Q6W8T4	Q6w8t4 euplotes mi
53	188	3.2	960	2	Q7R365	Q7r365 giardia lam
54	187	3.1	260	2	Q6SXU6	Q6sxu6 euplotes ra
55	184.5	3.1	960	2	Q9NCP5	Q9ncp5 giardia lam
56	179	3.0	2675	2	Q9FB23	Q9fb23 streptomyce
57	175.5	2.9	3530	1	MY15_HUMAN	Q9ukn7 homo sapien
58	172.5	2.9	660	1	YHL1_EBV	P03181 epstein-bar
59	172.5	2.9	660	2	Q777A3	Q777a3 human herpe
60	166	2.8	1474	2	Q962F9	Q962f9 cryptospori
61	161	2.7	1070	2	Q96JG5	Q96jg5 homo sapien
62	159	2.7	771	2	Q9XHY4	Q9xhy4 oryza sativ
63	158	2.6	850	2	Q66HA1	Q66ha1 rattus norv
64	154	2.6	243	2	Q6W8T3	Q6w8t3 euplotes eu
65	152.5	2.6	296	2	Q69118	Q69118 human herpe
66	151.5	2.5	669	2	Q8N4X0	Q8n4x0 homo sapien
67	150.5	2.5	653	1	SF01_MOUSE	Q64213 mus musculu
68	150	2.5	611	2	Q8K0M8	Q8k0m8 mus musculu
69	150	2.5	850	2	Q80X16	Q80x16 mus musculu
70	150	2.5	850	2	Q9J315	Q9j315 mus musculu
71	149.5	2.5	886	2	Q9VDK9	Q9vdk9 mus musculu
72	148.5	2.5	1064	2	Q6P7W6	Q6p7w6 mus musculu
73	148.5	2.5	1305	2	Q80YF9	Q80yf9 mus musculu
74	148	2.5	243	2	Q6SXU3	Q6sxu3 euplotes eu
75	147.5	2.5	986	2	Q9DW99	Q9dw99 rat cytomeg
76	147	2.5	1003	1	MBD6_HUMAN	Q96dn6 homo sapien
77	147	2.5	1003	2	Q6P0P0	Q6p0p0 homo sapien
78	147	2.5	1014	2	Q6W5R1	Q6w5r1 streptomyce
79	147	2.5	1186	2	Q7T401	Q7t401 cercopithe
80	146	2.4	442	2	Q9FCJ5	Q9fcj5 streptomyce
81	146	2.4	539	2	Q9SR71	Q9sr71 arabidopsis
82	146	2.4	953	2	Q91VT4	Q91vt4 homo sapien
83	146	2.4	1045	2	Q8N5P7	Q8n5p7 homo sapien
84	145	2.4	700	2	Q7T400	Q7t400 cercopithe
85	145	2.4	1074	2	Q6AT15	Q6ati5 oryza sativ
86	144	2.4	1523	2	Q93H10	Q93h10 streptomyce
87	143	2.4	1070	2	Q9APM8	Q9apm8 myxococcus
88	142.5	2.4	892	2	Q71JB1	Q71jb1 homo sapien
89	142.5	2.4	1560	2	Q60275	Q60275 homo sapien
90	142	2.4	1461	1	IE18_PRTVF	P11675 pseudorabie
91	141.5	2.4	889	2	Q9F2N5	Q9f2n5 streptomyce
92	141.5	2.4	950	2	Q82QX8	Q82qx8 streptomyce
93	141	2.4	501	2	Q8TEE9	Q8tee9 homo sapien
94	140.5	2.4	1892	2	Q97007	Q97007 leishmania
95	140	2.3	542	2	Q68872	Q68872 myxococcus
96	140	2.3	3493	2	Q7TMR6	Q7tmr6 mus musculu
97	140	2.3	3511	1	MY15_MOUSE	Q9qz24 mus musculu
98	139.5	2.3	190	2	Q6W8T5	Q6w8t5 euplotes va
99	139.5	2.3	552	2	Q9RSH9	Q9rsh9 deinococcus
100	139	2.3	425	2	O52231	O52231 streptococc

```

RESULT 1
TERT_HUMAN
ID TERT_HUMAN STANDARD; PRT; 1132 AA.
AC O14746; O14783;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic
DE subunit) (HST2) (Telomerase-associated protein 2) (TP2).
GN Name=TERT; Synonyms=EST2, TCSI, TRT;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97400623; PubMed=9252327; DOI=10.1126/science.277.5328.955;
RX Nakamura T.M., Morin G.B., Chapman K.B., Weinrich S.L., Andrews W.H.,
RX Lingner J., Harley C.B., Cech T.R.;
RA "Telomerase catalytic subunit homologs from fission yeast and human.";
RL Science 277:955-959(1997).
[2]
RC SEQUENCE FROM N.A.
RX MEDLINE=97433088; PubMed=9288757; DOI=10.1016/S0092-8674(00)80538-3;
RX Meyerson M., Counter C.M., Eaton E.N., Ellison L.W., Steiner P.,
RX Caddle S.D., Ziaugra L., Beijersbergen R.L., Davidoff M.J., Liu Q.,
RX Bacchetti S., Haber D.A., Weinberg R.A.;
RA "hEST2, the putative human telomerase catalytic subunit gene, is up-
RA regulated in tumor cells and during immortalization.";
RL Cell 90:785-795(1997).
[3]
RC SEQUENCE FROM N.A.
RX MEDLINE=99267414; PubMed=10333526; DOI=10.1016/S0378-1119(99)00108-0;
RX Wick M., Zubov D., Hagen G.;
RA "Genomic organization and promoter characterization of the gene
RA encoding the human telomerase reverse transcriptase (hTERT).";
RL Gene 232:97-106(1999).
[4]
RC SEQUENCE FROM N.A.
RX Londono-Vallejo J.A.;
RA "Sequence of a BAC carrying the entire hTERT gene.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[5]
RC FUNCTION IN TELOMERASE ACTIVITY, TISSUE SPECIFICITY, ASSOCIATION WITH
RC TEP1, AND MUTAGENESIS OF ASP-868; ASP-869 AND ASP-712.
RX PubMed=9389643;
RX Harrington L., Zhou W., McPhail T., Oulton R., Yeung D.S., Mar V.,
RX Baas M.B., Robinson M.O.;
RA "Human telomerase contains evolutionarily conserved catalytic and
RA structural subunits.";
RL Genes Dev. 11:3109-3115(1997).
[6]
RC ASSOCIATION WITH TEP1.
RX PubMed=11029039;
RX Beattie T.L., Zhou W., Robinson M.O., Harrington L.;
RA "Polymerization defects within human telomerase are distinct from
RA telomerase RNA and TEP1 binding.";
RL Mol. Biol. Cell 11:3329-3340(2000).
CC -!- FUNCTION: Telomerase is a ribonucleoprotein enzyme essential for
CC the replication of chromosome termini in most eukaryotes. It
CC elongates telomeres. It is a reverse transcriptase that adds
CC simple sequence repeats to chromosome ends by copying a template
CC sequence within the RNA component of the enzyme.
CC -!- SUBUNIT: Component of the telomerase ribonucleoprotein complex at
CC least composed of TEP1, EST1A, POT1 and a telomerase RNA template
CC component (TER). Interacts with PINK1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DISEASE: Activation of telomerase has been implicated in cell
CC immortalization and cancer cell pathogenesis.
CC -!- SIMILARITY: Belongs to the reverse transcriptase family.
CC Telomerase subfamily.

```

```

CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AF015950; AAC51672.1; -
CC EMBL; AF018167; AAC51724.1; -
CC EMBL; AF128894; AAD30037.1; -
CC EMBL; AF128893; AAD30037.1; JOINED.
CC EMBL; AY007685; AAG23289.1; -
CC F1R; T03844; T03844.
CC Genew; HGNC:11730; TERT.
CC MIM; 187270; -
CC GO; GO:0000781; C:Chromosome, telomeric region; TAS.
CC GO; GO:0042162; F:telomeric DNA binding; TAS.
CC GO; GO:0003721; F:telomeric template RNA reverse transcriptase. . .; TAS.
CC InterPro; IPR000477; RTase.
CC InterPro; IPR003545; Telomerase_RT.
CC Pfam; PF00078; RVT; 1.
CC PRINTS; PR01365; TELOMERASERT.
CC PROSITE; PS00878; RT_POL; 1.
CC DNA-binding; Nuclear protein; Ribonucleoprotein;
CC RNA-directed DNA polymerase; Telomere; Transcriptase.
CC DOMAIN 605 935
CC MUTAGEN 868 868 D->A: Loss of telomerase activity.
CC MUTAGEN 868 869 DD->AA: Loss of telomerase activity.
CC MUTAGEN 869 869 D->A: Loss of telomerase activity.
CC MUTAGEN 712 712 D->A: Loss of telomerase activity.
CC CONFLICT 516 516 D -> G (in Ref. 2).
CC SEQUENCE 1132 AA; 126996 MW; 94E35469C4CA33A0 CRC64;
Query Match 99.8%; Score 5952; DB 1; Length 1132;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MPRAPRCRAVRLSLRSHYREVLPLATFVRLRGLPGQWRLVQGDPAAPRALVAQCVCVPW 60
DB 1 MPRAPRCRAVRLSLRSHYREVLPLATFVRLRGLPGQWRLVQGDPAAPRALVAQCVCVPW 60
QY 61 DARPPPAAPSPFQVSCIKELVARVQLRCLRGAKNVLAFCGALLDGGARGGPEFTTSVR 120
DB 61 DARPPPAAPSPFQVSCIKELVARVQLRCLRGAKNVLAFCGALLDGGARGGPEFTTSVR 120
QY 121 SYLPTNTVDALRGSGAWGLLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
DB 121 SYLPTNTVDALRGSGAWGLLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
QY 181 ATQARPPPHASGPRRRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSASRLPLPKPRR 240
DB 181 ATQARPPPHASGPRRRRLGRCERAWNHSVREAGVPLGLPAPGARRRGGSASRLPLPKPRR 240
QY 241 GAAPERTPVGGQSWAHQRTGSDRGFCVVSPPARPAEATSEALSGSTRHSFVS 300
DB 241 GAAPERTPVGGQSWAHQRTGSDRGFCVVSPPARPAEATSEALSGSTRHSFVS 300
QY 301 RQHHAGPSTSRPRPMDTTPCPVYATKHFLYSSGDKQLRPSFLSSRLPSLTGARRL 360
DB 301 RQHHAGPSTSRPRPMDTTPCPVYATKHFLYSSGDKQLRPSFLSSRLPSLTGARRL 360
QY 361 VETIFLGSREPMFGTPRRRLPRLPQRYWQMRPLFLELGNHACPCYGVLLKTHCPLRAAVT 420
DB 361 VETIFLGSREPMFGTPRRRLPRLPQRYWQMRPLFLELGNHACPCYGVLLKTHCPLRAAVT 420
QY 421 PAAGVCAREXPQGSVAAPSEEDTPRLVQLLRQHSFPWQYGFVACLRRLVPPGLWGS 480
DB 421 PAAGVCAREXPQGSVAAPSEEDTPRLVQLLRQHSFPWQYGFVACLRRLVPPGLWGS 480
QY 481 RHNERRFLRNTKKFISLGKHALSLQLBTWKMSVRDCAWLRRSPGVGCVPAAEHRLREI 540

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 09:48:30 ; Search time 25 Seconds
(without alignments)
4356.695 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRSLLRSYRE.....TALEAANPALPSPDKYILD 1132

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5952	99.8	1132	2 T03844	telomerase catalyt
2	717.5	12.0	1123	2 T51517	telomerase reverse
3	588.5	9.9	989	2 T03838	telomerase catalyt
4	388	6.5	1132	2 T31107	telomerase reverse
5	357	6.0	1117	2 T14891	telomerase (SC 2.7
6	355.5	6.0	884	2 S53396	telomerase catalyt
7	175.5	2.9	3530	2 A59266	unconventional myo
8	172.5	2.9	660	1 Q0BE3	BHLF1 protein - hu
9	142.5	2.4	1560	2 T00080	hypothetical prote
10	140.5	2.4	1892	2 T18314	hypothetical prote
11	140	2.3	1460	1 EDBE1F	immediate-early pr
12	140	2.3	3511	2 A59295	unconventional myo
13	139.5	2.3	552	2 F75311	ABC transporter, A
14	138.5	2.3	924	2 S27923	gene LF3 protein -
15	135	2.3	1356	1 C45219	N-methyl-D-asparta
16	134	2.2	606	2 G75302	orotidine 5'-phosp
17	133	2.2	1106	2 J00405	hypothetical 119.5
18	132.5	2.2	946	2 J07810	inositol 1,4,5-tri
19	131.5	2.2	1184	2 G01763	atrophin-1 - human
20	131	2.2	916	2 J070396	reverse transcript
21	131	2.2	1446	1 A45344	immediate-early pr
22	130.5	2.2	1184	2 S08032	atrophin-1 - human
23	130.5	2.2	2715	2 T13049	eyelid - fruit fly
24	130	2.2	1067	2 T18196	pol protein - silk
25	129.5	2.2	383	2 S32975	gene BCREF2 protein
26	129	2.2	403	2 S22796	PRP2 protein - hu
27	129	2.2	860	2 S55543	RNA-directed DNA p
28	128.5	2.2	628	2 S01955	hypothetical prote
29	128	2.1	376	2 C75580	adenine deaminase-

30	128	2.1	1776	2 G86280	protein T5E21.13 (
31	127.5	2.1	1048	2 T31425	C-terminal domain-
32	127	2.1	260	2 S22373	proline-rich prote
33	127	2.1	505	2 S72273	actin-depolymerizi
34	127	2.1	580	2 T43481	probable mucin DXF
35	127	2.1	847	1 A53800	mixed-lineage prot
36	127	2.1	862	2 T46289	hypothetical prote
37	127	2.1	1039	2 T35878	hypothetical prote
38	126.5	2.1	603	2 H75272	probable nucleic a
39	126	2.1	330	2 E98119	transposase, uncha
40	125.5	2.1	574	2 T43556	Wiskott-Aldrich sy
41	125.5	2.1	574	2 T38819	viral proteinase -
42	125	2.1	522	2 S52216	tenascin Y precurs
43	125	2.1	1914	2 T42635	immediate-early pr
44	124	2.1	1298	1 EDBE75	N-methyl-D-asparta
45	124	2.1	1323	2 S27224	hypothetical prote
46	123.5	2.1	381	2 S16506	synapsin Ia - rat
47	123.5	2.1	704	2 A30411	68.6K capsid prote
48	123	2.1	646	1 W2BEC8	salivary proline-r
49	122.5	2.1	310	1 P1HUSD	hypothetical prote
50	122.5	2.1	635	2 F75477	hypothetical prote
51	122.5	2.1	915	2 T12526	otogelin - mouse
52	122.5	2.1	2910	2 T42214	MHC class III hist
53	122	2.0	1870	2 S37671	hydroxyproline-ric
54	121.5	2.0	620	2 S06733	guanylate cyclase
55	121.5	2.0	1102	2 JH0717	period protein PER
56	121.5	2.0	1257	2 T13957	MHC class III hist
57	121	2.0	2142	2 B35098	licheninase (EC 3.
58	120.5	2.0	335	2 T05722	probable proline-r
59	120.5	2.0	891	2 G84693	ataxin-2 - mouse
60	120.5	2.0	1285	2 T14171	probable membrane
61	120	2.0	509	2 T34871	BOLFI protein - hu
62	120	2.0	1239	1 Q0BE10	DNA-directed RNA p
63	120	2.0	1859	1 A34092	DNA-directed RNA p
64	120	2.0	1862	2 T29959	MHC class III hist
65	120	2.0	1872	2 S36152	hypothetical prote
66	120	2.0	2793	2 B90784	hypothetical prote
67	120	2.0	2806	2 D85644	hypothetical prote
68	119.5	2.0	361	2 T12543	immediate-early pr
69	119.5	2.0	775	1 EDBE11	transcription fact
70	119	2.0	639	2 G02919	FK506 polyketide s
71	119	2.0	7576	2 T17428	hypothetical prote
72	118.5	2.0	628	2 S19150	hypothetical prote
73	118.5	2.0	751	2 D98320	N-Chimerin homolog
74	118.5	2.0	903	2 T00705	probable multi-dom
75	118.5	2.0	1334	2 T50568	fatty-acid synthas
76	118.5	2.0	2796	2 J04743	UL36 protein - hum
77	118.5	2.0	3164	1 WMBEH6	CWI7R protein - mo
78	118	2.0	548	2 S52735	probable non-ribos
79	118	2.0	2352	2 C83229	hypothetical prote
80	117.5	2.0	312	2 A61183	probable serine/th
81	117.5	2.0	576	2 T36729	hypothetical prote
82	117.5	2.0	1222	2 T22490	collagen alpha 1(V
83	117.5	2.0	2944	2 A54849	proline-rich prote
84	117	2.0	300	2 S19560	F02569.2 protein (
85	117	2.0	512	2 E59437	probable integral
86	117	2.0	684	2 T36771	synapsin I - rat
87	117	2.0	691	2 A25704	latrophilin-1, bra
88	117	2.0	1467	2 T18411	latrophilin-1, bra
89	117	2.0	1472	2 T18413	UL43 protein - hum
90	116.5	2.0	434	1 WMBEF3	probable oxidoredu
91	116.5	2.0	1174	2 B85740	probable oxidoredu
92	116.5	2.0	1174	2 H90878	hairless protein -
93	115.5	1.9	1182	2 I48378	hypothetical prote
94	115.5	1.9	1322	2 T24140	proline-rich prote
95	115	1.9	301	2 E29149	proline-rich prote
96	115	1.9	309	2 S10889	50kD proline rich
97	115	1.9	456	2 T35474	membrane transloca
98	115	1.9	629	2 E47096	probable DNA-bindi
99	115	1.9	753	2 G87178	DNA-binding protei
100	115	1.9	767	2 S41479	

ALIGNMENTS

RESULT 1
T03844
telomerase catalytic chain - human
N;Alternate names: telomerase reverse transcriptase
C;Species: Homo sapiens (man)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03844
R;Nakamura, T.M.; Morin, G.B.; Chapman, K.B.; Weinrich, S.L.; Andrews, W.H.; Lingner, J.
Science 277, 955-959, 1997
A;Title: Telomerase catalytic subunit homologs from fission yeast and human.
A;Reference number: Z15111; MUID:97400623; PMID:9252327
A;Accession: T03844
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1132 <NAK>
A;Cross-references: UNIPROT:O14746; EMBL:AF015950; NID:g23330016; PIDN:AAC51672.1; PID:g23330016
A;Experimental source: kidney
C;Genetics:
A;Gene: TRT
A;Map position: 5p

Query Match 99.8%; Score 5952; DB 2; Length 1132;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLLGPGQWRVLRQGDPAAPRALVAQCLVCPW 60
Db 1 MPAPRCRAVRSLLRSHYREVLPATFVRRLLGPGQWRVLRQGDPAAPRALVAQCLVCPW 60

Qy 61 DARPPAPSPROVSCLELVARVLORLCERGAKNVAFGALLDARGGPEAFTTSVR 120
Db 61 DARPPAPSPROVSCLELVARVLORLCERGAKNVAFGALLDARGGPEAFTTSVR 120

Qy 121 SYLPTNTVDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Db 121 SYLPTNTVDALRGSGAWGLLRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180

Qy 181 ATQARPPPHAGSPRRRLCERAWNHSVREAGVPLGAPAGARRGGASRSRLPKRPRR 240
Db 181 ATQARPPPHAGSPRRRLCERAWNHSVREAGVPLGAPAGARRGGASRSRLPKRPRR 240

Qy 241 GAAPEPERTVGGQSWAHPGTRGSDRGFCVWSPARPAEATSLGALSSTRHSPVG 300
Db 241 GAAPEPERTVGGQSWAHPGTRGSDRGFCVWSPARPAEATSLGALSSTRHSPVG 300

Qy 301 ROHAGPPTSRPPRPWDTPCPPVVAETKHFYSGDKQLRPGFLLSSLRPSLTGARRL 360
Db 301 ROHAGPPTSRPPRPWDTPCPPVVAETKHFYSGDKQLRPGFLLSSLRPSLTGARRL 360

Qy 361 VETIFLGSRPWMPGTPRRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 420
Db 361 VETIFLGSRPWMPGTPRRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPRAAVT 420

Qy 421 PAAGVCAREKQGSVAAPBEEDTDPRRLVQLLRQHSSPWQYGFVRACLRRLVPPGLWGS 480
Db 421 PAAGVCAREKQGSVAAPBEEDTDPRRLVQLLRQHSSPWQYGFVRACLRRLVPPGLWGS 480

Qy 481 RHNERFRNTKFTSLGKHAKLSIQELTWKMSVRDCAWLRRSPGVCPVPAEHRLEEI 540
Db 481 RHNERFRNTKFTSLGKHAKLSIQELTWKMSVRDCAWLRRSPGVCPVPAEHRLEEI 540

Qy 541 LAKFLHLMVSVVVELLSFFVTTTTFQKNLFFRPSVMSKLSQIGIRQHLKRVQLRE 600
Db 541 LAKFLHLMVSVVVELLSFFVTTTTFQKNLFFRPSVMSKLSQIGIRQHLKRVQLRE 600

Qy 601 LSEAEVRQHRARPALLTSRLRFPKPDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660
Db 601 LSEAEVRQHRARPALLTSRLRFPKPDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA 660

Qy 661 LFSVLNYERARRPGLGASVLGLDDIHRAWRTFVLVRQAQDPPPELYFVKVDVTGAYDTI 720

Db 661 LFSVLNYERARRPGLGASVLGLDDIHRAWRTFVLVRQAQDPPPELYFVKVDVTGAYDTI 720

Qy 721 PQDLTEVIASIIKPONTYCVRRYAVVQKAAGHVRKAFKSHVSTLTDLPYMQFVAHL 780
Db 721 PQDLTEVIASIIKPONTYCVRRYAVVQKAAGHVRKAFKSHVSTLTDLPYMQFVAHL 780

Qy 781 QETSPRLDADVIEOSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVOCQGIPOQSILSTL 840
Db 781 QETSPRLDADVIEOSSSLNEASSGLFDVFLRFMCHHAVIRGKSYVOCQGIPOQSILSTL 840

Qy 841 LCSLCYGDMMENKLPAGIRRDGLLLRLVDDFLVTPHLLTHAKTFURTLVRGVPYGCVVNL 900
Db 841 LCSLCYGDMMENKLPAGIRRDGLLLRLVDDFLVTPHLLTHAKTFURTLVRGVPYGCVVNL 900

Qy 901 RKTVNFPEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVDQSDYSYARTSTRASVTF 960
Db 901 RKTVNFPEDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVDQSDYSYARTSTRASVTF 960

Qy 961 NRGFKAGNRMRKLFGLRLKCHSLFDLQVNSLQTVCTNIYKILLQAVRFHACVQLQP 1020
Db 961 NRGFKAGNRMRKLFGLRLKCHSLFDLQVNSLQTVCTNIYKILLQAVRFHACVQLQP 1020

Qy 1021 FHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAPLL 1080
Db 1021 FHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSLGAKGAAGPLPSEAVQWLCHQAPLL 1080

Qy 1081 KLTRHRVTVYVPLGSLRTAQTLRSKLPGTTLTALEAAANPALPSDFKTILD 1132
Db 1081 KLTRHRVTVYVPLGSLRTAQTLRSKLPGTTLTALEAAANPALPSDFKTILD 1132

RESULT 2
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N;Alternate names: protein F5E19_190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51517
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1123 <SAT>
A;Cross-references: UNIPROT:Q9SPU7; EMBL:AL391147
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2
A;Note: F5E19_190

Query Match 12.0%; Score 717.5; DB 2; Length 1123;
Best Local Similarity 23.7%; Pred. No. 9e-41;
Matches 293; Conservative 199; Mismatches 512; Indels 233; Gaps 39;

Qy 1 MPAPRCRAVRSLLRSHYREVLPATFV-----RRLGPGQWR-----L 38
Db 1 MPAPRCRAVRSLLRSHYREVLPATFV-----RRLGPGQWR-----L 38

Qy 39 VQCGDPAAPRALVAQCLVCPWDARPPPAAPSPRQVSCLELVARVLQR-----CERGAKN 95
Db 39 VQCGDPAAPRALVAQCLVCPWDARPPPAAPSPRQVSCLELVARVLQR-----CERGAKN 95

Qy 61 LRSDDPPIHYRKLHLHRCFV-VLHEQTPELDFSPTSWWSQREIVREIEMMQSGCD--CQN 117
Db 61 LRSDDPPIHYRKLHLHRCFV-VLHEQTPELDFSPTSWWSQREIVREIEMMQSGCD--CQN 117

Qy 96 VLAFGFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLAR 155
Db 96 VLAFGFALLDARGGPPPEAFTTSVRSYLPNTVTDALRGSGAWGLLRVGGDDVLVHLLAR 155

Qy 118 VICARYDKYDQS-----SPILLELT-SSSWBEFLAKRVGHDVWVLLQQ 159
Db 118 VICARYDKYDQS-----SPILLELT-SSSWBEFLAKRVGHDVWVLLQQ 159

Qy 156 CALFVLVAPSCAYQVCGPPLYQLGAAT-----QARPPPHASGPRRLRCERAWNHSVRE 209
Db 156 CALFVLVAPSCAYQVCGPPLYQLGAAT-----QARPPPHASGPRRLRCERAWNHSVRE 209

Qy 160 TSIFPLPLGLKKHQVSGPPLCIKHRTLSVHENKRKRDNDVQPPTKRQWLSSAVDDCPKD 219
Db 160 TSIFPLPLGLKKHQVSGPPLCIKHRTLSVHENKRKRDNDVQPPTKRQWLSSAVDDCPKD 219

Index	Value	Label
1	5963	Sequence 344, App
2	5963	Sequence 344, App
3	5963	Sequence 344, App
4	5963	Sequence 344, App
5	5963	Sequence 334, App
6	5963	Sequence 334, App
7	5963	Sequence 628, App
8	5963	Sequence 628, App
9	5963	Sequence 628, App
10	5963	Sequence 628, App
11	5952	Sequence 2, Appli

85 3496 58.6 1122 16 US-10-602-441-4 Sequence 4, Appli
86 3496 58.6 1122 16 US-10-862-698-2 Sequence 2, Appli
87 3496 58.6 1122 17 US-10-794-514A-5 Sequence 5, Appli
88 2125 52.4 622 14 US-10-294-778-12 Sequence 12, Appli
89 2581 43.3 500 14 US-10-282-960-81 Sequence 81, Appli
90 2253 37.8 1131 14 US-10-295-681-52 Sequence 52, Appli
91 2178 36.5 438 14 US-10-294-778-10 Sequence 10, Appli
92 2167.5 36.3 743 16 US-10-602-441-10 Sequence 10, Appli
93 2163.5 36.3 437 14 US-10-294-778-2 Sequence 2, Appli
94 2079 34.9 564 9 US-08-843-676-101 Sequence 101, App
95 2079 34.9 564 9 US-09-766-253-101 Sequence 101, App
96 2079 34.9 564 10 US-09-438-486-101 Sequence 101, App
97 2079 34.9 564 14 US-10-053-758-101 Sequence 101, App
98 2079 34.9 564 14 US-10-054-295-101 Sequence 101, App
99 2079 34.9 564 14 US-10-054-611-101 Sequence 101, App
100 2079 34.9 564 15 US-10-325-810-267 Sequence 267, App

ALIGNMENTS

RESULT 1
US-10-325-810-344
; Sequence 344, Application US/10325810
; Publication No. US20030204069A1
; GENERAL INFORMATION:
; APPLICANT: Gech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin B.
; Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 633
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/325,810
; FILING DATE: 20-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,181
; FILING DATE: 29-Sep-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:

NAME: Ausenhuss, Scott L.
REGISTRATION NUMBER: 42,271
REFERENCE/DOCKET NUMBER: 015389-002620US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 344:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 344:
US-10-325-810-344
Query Match 100.0%; Score 5963; DB 15; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPRAPRCRAVRSLLRSHYREVLPPLATEFVRLPGQWRLVORGDPAAAFRALVAQCLVCVPW 60
Db 1 MPRAPRCRAVRSLLRSHYREVLPPLATEFVRLPGQWRLVORGDPAAAFRALVAQCLVCVPW 60
Qy 61 DARPPPAAPSFQVSCLELVARVQLRCLERGAKNVLAFCGAFALLDGARGGPPPEAFTTSVR 120
Db 61 DARPPPAAPSFQVSCLELVARVQLRCLERGAKNVLAFCGAFALLDGARGGPPPEAFTTSVR 120
Qy 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Db 121 SYLPNTVTDALRGSGAWGLLLRRVGGDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Qy 181 ATQARPPPHASGPRRLRGCEANWHSVRBAGVPLGLPAPGARRRGGSASLSLPKPRR 240
Db 181 ATQARPPPHASGPRRLRGCEANWHSVRBAGVPLGLPAPGARRRGGSASLSLPKPRR 240
Qy 241 GAAPEPERTVGGQSWAHPCGTRGSDRGFCVVSAPPAEAEATSEALSGSTRHSFVS 300
Db 241 GAAPEPERTVGGQSWAHPCGTRGSDRGFCVVSAPPAEAEATSEALSGSTRHSFVS 300
Qy 301 RQHHAGPSTSRPPRWDTPCPVYATKIFLVSQDKQEOIRPSFLSSRLPSLTGARRL 360
Db 301 RQHHAGPSTSRPPRWDTPCPVYATKIFLVSQDKQEOIRPSFLSSRLPSLTGARRL 360
Qy 361 VETIFLGSRPMPGCTPRLPRLPQRYQWMPPLFLLELGNHAQCPYGVLLKTHCPLRAVT 420
Db 361 VETIFLGSRPMPGCTPRLPRLPQRYQWMPPLFLLELGNHAQCPYGVLLKTHCPLRAVT 420
Qy 421 PAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRQHSPPWQVYGFVRACLRLVPPGLWGS 480
Db 421 PAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRQHSPPWQVYGFVRACLRLVPPGLWGS 480
Qy 481 RHNERRLNTKKFISLGKHAQLSLQBLTWMSVRDCAWLRSPGVCVCPAAEHLREEI 540
Db 481 RHNERRLNTKKFISLGKHAQLSLQBLTWMSVRDCAWLRSPGVCVCPAAEHLREEI 540
Qy 541 LAKFLHMLSVVYVELLSRPFYVTTETTFQKNRLFYRPSVMSKLQSIGIRKHLKRVOLRE 600
Db 541 LAKFLHMLSVVYVELLSRPFYVTTETTFQKNRLFYRPSVMSKLQSIGIRKHLKRVOLRE 600
Qy 601 LSEAEVQHQREARPAALITSRLRFPKPDGLRPIVNMDDYVVGARTFRREKRAERLTSRKA 660
Db 601 LSEAEVQHQREARPAALITSRLRFPKPDGLRPIVNMDDYVVGARTFRREKRAERLTSRKA 660
Qy 661 LFSVLNYERARRPGLLGASVLGLDDIHRARTFVLRVRAQDPPPELFFVKVDVTGAYDTI 720
Db 661 LFSVLNYERARRPGLLGASVLGLDDIHRARTFVLRVRAQDPPPELFFVKVDVTGAYDTI 720
Qy 721 PQDLTEVIASIIKPNQTYCVRRYAVVQKAAHGVKAFKSHVSTLLDLOPYMQFVAHL 780
Db 721 PQDLTEVIASIIKPNQTYCVRRYAVVQKAAHGVKAFKSHVSTLLDLOPYMQFVAHL 780
Qy 781 QETSPLRDADVIEQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPOGSIISLTL 840
Db 781 QETSPLRDADVIEQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPOGSIISLTL 840

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 10:42:15 ; Search time 27 Seconds
(without alignments)
3129.731 Million cell updates/sec

Title: US-08-974-584C-118
Perfect score: 5963
Sequence: 1 MPRAPRCRANRSLRSHYRE.....TALEAANPALPSPDKTILD 1132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/ECTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5963	100.0	1132	3	US-08-974-549A-344
2	5963	100.0	1132	4	US-09-402-181B-344
3	5963	100.0	1132	4	US-09-721-456-344
4	5963	100.0	1407	3	US-08-974-549A-628
5	5963	100.0	1407	4	US-08-912-951-334
6	5963	100.0	1407	4	US-09-402-181B-628
7	5963	100.0	1407	4	US-09-721-456-628
8	5952	99.8	1132	3	US-08-851-843A-225
9	5952	99.8	1132	3	US-08-974-549A-2
10	5952	99.8	1132	3	US-08-854-050-225
11	5952	99.8	1132	3	US-09-430-323-225
12	5952	99.8	1132	3	US-09-128-354-2
13	5952	99.8	1132	4	US-09-675-321-2
14	5952	99.8	1132	4	US-09-052-919-2
15	5952	99.8	1132	4	US-08-912-951-2
16	5952	99.8	1132	4	US-09-402-181B-2
17	5952	99.8	1132	4	US-09-721-456-2
18	5952	99.8	1132	4	US-09-953-052-2
19	5952	99.8	1132	4	US-09-042-460-3
20	5952	99.8	1132	4	US-09-949-016-6326
21	5952	99.8	1154	3	US-08-974-549A-611
22	5952	99.8	1154	3	US-08-912-951-323
23	5952	99.8	1154	4	US-09-402-181B-611
24	5952	99.8	1154	4	US-09-721-456-611
25	5952	99.8	1189	3	US-08-974-549A-613
26	5952	99.8	1189	4	US-08-912-951-325
27	5952	99.8	1189	4	US-09-402-181B-613

28	5952	99.8	1189	4	US-09-721-456-613	Sequence 613, App
29	5952	99.8	1200	3	US-08-974-549A-612	Sequence 614, App
30	5952	99.8	1200	4	US-08-912-951-324	Sequence 615, App
31	5952	99.8	1200	4	US-09-402-181B-612	Sequence 616, App
32	5952	99.8	1200	4	US-09-721-456-612	Sequence 617, App
33	5952	99.8	1285	3	US-08-974-549A-600	Sequence 600, App
34	5952	99.8	1285	4	US-08-912-951-314	Sequence 601, App
35	5952	99.8	1285	4	US-09-402-181B-600	Sequence 602, App
36	5952	99.8	1285	4	US-09-721-456-600	Sequence 603, App
37	5798	97.2	1193	4	US-09-949-016-11712	Sequence 11712, A
38	4046	67.9	807	3	US-08-974-549A-5	Sequence 5, Appli
39	4046	67.9	807	4	US-08-912-951-5	Sequence 5, Appli
40	4046	67.9	807	4	US-09-402-181B-5	Sequence 5, Appli
41	4046	67.9	807	4	US-09-721-456-5	Sequence 5, Appli
42	3810.5	63.9	1003	3	US-08-851-843A-217	Sequence 217, App
43	3810.5	63.9	1003	3	US-08-974-549A-336	Sequence 336, App
44	3810.5	63.9	1003	3	US-08-854-050-217	Sequence 217, App
45	3810.5	63.9	1003	3	US-09-430-323-217	Sequence 217, App
46	3810.5	63.9	1003	3	US-09-402-181B-336	Sequence 336, App
47	3810.5	63.9	1003	4	US-09-721-456-336	Sequence 336, App
48	3496	58.6	1122	4	US-09-042-460-2	Sequence 2, Appli
49	3125	52.4	622	4	US-09-582-924B-12	Sequence 12, Appli
50	2178	36.5	438	4	US-09-582-924B-10	Sequence 10, Appli
51	2163.5	36.3	437	4	US-09-582-924B-2	Sequence 2, Appli
52	2079	34.9	564	3	US-08-851-843A-101	Sequence 101, App
53	2079	34.9	564	3	US-08-974-549A-267	Sequence 267, App
54	2079	34.9	564	3	US-08-854-050-101	Sequence 101, App
55	2079	34.9	564	3	US-09-430-323-101	Sequence 101, App
56	2079	34.9	564	4	US-09-402-181B-267	Sequence 267, App
57	2079	34.9	564	4	US-09-721-456-267	Sequence 267, App
58	2079	34.9	564	4	US-09-766-253-101	Sequence 101, App
59	1666.5	27.9	364	4	US-09-417-485D-40	Sequence 40, Appli
60	1565	26.2	538	3	US-08-974-549A-602	Sequence 602, App
61	1565	26.2	538	4	US-08-912-951-316	Sequence 316, App
62	1565	26.2	538	4	US-09-402-181B-602	Sequence 602, App
63	1565	26.2	538	4	US-09-721-456-602	Sequence 602, App
64	1506	25.3	514	3	US-08-974-549A-605	Sequence 605, App
65	1506	25.3	514	4	US-08-912-951-319	Sequence 319, App
66	1506	25.3	514	4	US-09-402-181B-605	Sequence 605, App
67	1506	25.3	514	4	US-09-721-456-605	Sequence 605, App
68	1454	24.4	517	3	US-08-974-549A-606	Sequence 606, App
69	1454	24.4	517	4	US-08-912-951-320	Sequence 320, App
70	1454	24.4	517	4	US-09-402-181B-606	Sequence 606, App
71	1454	24.4	517	4	US-09-721-456-606	Sequence 606, App
72	1447	24.3	530	3	US-08-974-549A-603	Sequence 603, App
73	1447	24.3	530	4	US-08-912-951-317	Sequence 317, App
74	1447	24.3	530	4	US-09-402-181B-603	Sequence 603, App
75	1447	24.3	530	4	US-09-721-456-603	Sequence 603, App
76	1438	24.1	283	4	US-09-424-226-6	Sequence 6, Appli
77	1386	23.2	515	3	US-08-974-549A-604	Sequence 604, App
78	1386	23.2	515	4	US-08-912-951-318	Sequence 318, App
79	1386	23.2	515	4	US-09-402-181B-604	Sequence 604, App
80	1386	23.2	515	4	US-09-721-456-604	Sequence 604, App
81	1243.5	20.9	364	4	US-09-417-485D-41	Sequence 41, Appli
82	1090	18.3	259	3	US-08-974-549A-10	Sequence 10, Appli
83	1090	18.3	259	4	US-08-912-951-10	Sequence 10, Appli
84	1090	18.3	259	4	US-09-402-181B-10	Sequence 10, Appli
85	1090	18.3	259	4	US-09-721-456-10	Sequence 10, Appli
86	929	15.6	222	3	US-08-851-843A-202	Sequence 202, App
87	929	15.6	222	3	US-08-974-549A-321	Sequence 321, App
88	929	15.6	222	3	US-08-854-050-202	Sequence 202, App
89	929	15.6	222	3	US-09-430-323-202	Sequence 202, App
90	929	15.6	222	4	US-09-402-181B-321	Sequence 321, App
91	929	15.6	222	4	US-09-721-456-321	Sequence 321, App
92	905.5	15.2	330	3	US-08-851-843A-203	Sequence 203, App
93	905.5	15.2	330	3	US-08-974-549A-322	Sequence 322, App
94	905.5	15.2	330	3	US-08-854-050-203	Sequence 203, App
95	905.5	15.2	330	3	US-09-430-323-203	Sequence 203, App
96	905.5	15.2	330	4	US-09-402-181B-322	Sequence 322, App
97	905.5	15.2	330	4	US-09-721-456-322	Sequence 322, App
98	728	12.2	381	4	US-09-417-485D-47	Sequence 47, Appli
99	646	10.8	129	3	US-08-851-843A-67	Sequence 67, Appli
100	646	10.8	129	3	US-08-974-549A-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1

US-08-974-549A-344
; Sequence 344, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 344:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1132 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-974-549A-344

Query Match 100.0%; Score 5963; DB 3; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MPRAPRCRAVRSLLRSHYREVLP	PLATFVRRLGPGQWRLVQGDPAAP	FRALVAQC	LCVCPW	60	
Db	1	MPRAPRCRAVRSLLRSHYREVLP	PLATFVRRLGPGQWRLVQGDPAAP	FRALVAQC	LCVCPW	60	
Qy	61	DARPPPAAPSFQVSCIKELVARV	QLCERGAQNVLAFGFALLDARGG	PPPEAFTT	SVR	120	
Db	61	DARPPPAAPSFQVSCIKELVARV	QLCERGAQNVLAFGFALLDARGG	PPPEAFTT	SVR	120	
Qy	121	SYLPNTVTDALRGSGAMGLLRV	GGDVLVHLLARCALFVLVAPSC	AYQVCGP	PLYQLGA	180	
Db	121	SYLPNTVTDALRGSGAMGLLRV	GGDVLVHLLARCALFVLVAPSC	AYQVCGP	PLYQLGA	180	
Qy	181	ATQARPPPHASGPRRLGRCERA	NHNSVREAGVPLGLPAPGARRR	GGASRSL	PLPKRRR	240	
Db	181	ATQARPPPHASGPRRLGRCERA	NHNSVREAGVPLGLPAPGARRR	GGASRSL	PLPKRRR	240	
Qy	241	GAAPERTPVGOGSWAHGRTGR	PSDRGFCVVSPPARPAEATSL	EGALSTR	HSHPSVG	300	
Db	241	GAAPERTPVGOGSWAHGRTGR	PSDRGFCVVSPPARPAEATSL	EGALSTR	HSHPSVG	300	
Qy	301	QOHAGPSTSRPPRPMWDTCP	PPVYAEKHFLYSSGDKQLRPS	FLSSLRPS	LTGARRL	360	
Db	301	QOHAGPSTSRPPRPMWDTCP	PPVYAEKHFLYSSGDKQLRPS	FLSSLRPS	LTGARRL	360	
Qy	361	VETIFLGSRPMPCTPRRLPRL	FORVQMRPLFLELGNHAQCPY	GYVLLKTHC	PLRAAVT	420	
Db	361	VETIFLGSRPMPCTPRRLPRL	FORVQMRPLFLELGNHAQCPY	GYVLLKTHC	PLRAAVT	420	
Qy	421	PAAGVCAREXPQGSVAAP	EEEDTPRRLVOLLQHSSPWQY	GFVACILRL	VPGLWGS	480	
Db	421	PAAGVCAREXPQGSVAAP	EEEDTPRRLVOLLQHSSPWQY	GFVACILRL	VPGLWGS	480	
Qy	481	RHNERFLRNTKKFISLGKHA	KLQELTWKMSVRDCAWLRSP	GVGCVPA	AAEHLREEI	540	
Db	481	RHNERFLRNTKKFISLGKHA	KLQELTWKMSVRDCAWLRSP	GVGCVPA	AAEHLREEI	540	
Qy	541	LAKPLHLMMSVYVVELLSR	SPFYVTETTFQKNRLFYRPS	VMSKLSQ	IGIRQHLKRV	600	
Db	541	LAKPLHLMMSVYVVELLSR	SPFYVTETTFQKNRLFYRPS	VMSKLSQ	IGIRQHLKRV	600	
Qy	601	LSEAEVROHREARPA	LLTSRLRIPKPDGLRPIVNDY	VVGARTFRE	KRAERLTSVKA	660	
Db	601	LSEAEVROHREARPA	LLTSRLRIPKPDGLRPIVNDY	VVGARTFRE	KRAERLTSVKA	660	
Qy	661	LFSVLNYERARRPGLLGAS	VLGLDDIHRAWRTFVLVRAO	DPPELPYFV	KVDVTGAYDTI	720	
Db	661	LFSVLNYERARRPGLLGAS	VLGLDDIHRAWRTFVLVRAO	DPPELPYFV	KVDVTGAYDTI	720	
Qy	721	PDRLTEVIASIIKPNQTY	CVRRYAVVQKAAHGVKAFK	SHVSTLTD	LPYMQFVAHL	780	
Db	721	PDRLTEVIASIIKPNQTY	CVRRYAVVQKAAHGVKAFK	SHVSTLTD	LPYMQFVAHL	780	
Qy	781	QETSPLRDAVVIEQSSSL	NEASSGLFDVFLRPMCHAVR	IRGKSVYV	QCQIGPQSILSTL	840	
Db	781	QETSPLRDAVVIEQSSSL	NEASSGLFDVFLRPMCHAVR	IRGKSVYV	QCQIGPQSILSTL	840	
Qy	841	LCSLCYGDMENKLFAGIR	RRDGLLLRLVDDFLVTPL	THAKTFLR	TLVRGVPEYGC	VVNL	900
Db	841	LCSLCYGDMENKLFAGIR	RRDGLLLRLVDDFLVTPL	THAKTFLR	TLVRGVPEYGC	VVNL	900
Qy	901	RKTVMNPFVEDEALGGT	AFVQMPAHGLFPWCGLLD	TRTLLEVQ	SDYSYARTS	IRASVTF	960
Db	901	RKTVMNPFVEDEALGGT	AFVQMPAHGLFPWCGLLD	TRTLLEVQ	SDYSYARTS	IRASVTF	960

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 02:52:05 ; Search time 80 Seconds
(without alignments)
5472.659 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963
Sequence: 1 MFPRAPRCRAVRSLLRSHYRE.....TALEAANPALPSPDKTILD 1132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5963	100.0	1132	2	AAW56113 Human tel
2	5952	99.8	1132	2	AAW46957 Human tel
3	5952	99.8	1132	2	AAW90251 Human cat
4	5952	99.8	1132	2	AAW28881 Human tel
5	5952	99.8	1132	2	AAW32090 Human tel
6	5952	99.8	1132	2	AAW43621 A human t
7	5952	99.8	1132	2	AAW26580 Human tel
8	5952	99.8	1132	4	AAW64859 Heart mus
9	5952	99.8	1132	4	AAW64329 Human pro
10	5952	99.8	1132	4	AAW99930 Human tel
11	5952	99.8	1132	4	AAW82765 Human tel
12	5952	99.8	1132	5	AAW29226 Human tel
13	5952	99.8	1132	5	AAW27235 Human tel
14	5952	99.8	1132	6	AAW42384 Human tel
15	5952	99.8	1132	6	AAW42063 Human tel
16	5952	99.8	1132	6	AAW56676 Human tel
17	5952	99.8	1132	6	AAW58045 Human tel
18	5952	99.8	1132	7	AAW21420 Human TER
19	5952	99.8	1132	7	AAW72743 ADH72743
20	5952	99.8	1132	8	AAW70114 hTERT pro
21	5952	99.8	1132	8	AAW90599 Human TER
22	5952	99.8	1132	8	AAW182172 Human tel
23	5952	99.8	1132	8	AAW70482 Human tel
24	5952	99.8	1154	2	AAW61350 Human tel
25	5952	99.8	1189	2	AAW47008 Glutathio

26	5946	99.7	1285	2	AAW47000
27	5945	99.7	1132	2	AAW71376 Human tel
28	5945	99.7	1132	2	AAW00627 Human tel
29	5945	99.7	1132	2	AAW00638 Truncated
30	5945	99.7	1132	2	AAW28401 Human EST
31	5945	99.7	1132	3	AAW96566 hEST2, a
32	5945	99.7	1132	7	ADC47061 Human TER
33	5945	99.7	1132	7	ADC40482 Human tel
34	5929	99.4	1405	2	AAW56101 Enhanced
35	5918	99.2	1166	2	AAW00647 Telomeras
36	5902.5	99.0	1199	2	AAW47007 Glutathio
37	5873	98.5	1120	2	AAW00641 Telomeras
38	5864	98.3	1120	2	AAW00650 Telomeras
39	5712	95.8	1150	2	AAW47006 Glutathio
40	5546	93.0	1053	2	AAW00640 Altered C
41	5507	92.4	1093	2	AAW00649 Altered C
42	5458	91.5	1041	2	AAW00652 Altered C
43	5458	91.5	1041	2	AAW00643 Altered C
44	5002	83.9	948	2	AAW00639 N-termina
45	4998	83.8	948	2	AAW00648 Truncated
46	4926	82.6	936	2	AAW00642 Truncated
47	4917	82.5	936	2	AAW00651 Truncated
48	4894	82.1	949	2	AAW61349 Human tel
49	4506	75.6	1152	8	ADG90609 TERT con
50	4046	67.9	807	2	AAW46997 Human tel
51	4046	67.9	807	2	AAW00637 N-termina
52	4046	67.9	807	2	AAW00646 Truncated
53	3619	60.7	1128	7	ADD21416 Golden ha
54	3619	60.7	1128	8	ADG90603 Hamster T
55	3496	58.6	1122	2	AAW28579 Murine TE
56	3496	58.6	1122	8	ADG90601 Murine TE
57	3466	58.1	1122	5	ABB06711 Mouse tel
58	3238	54.3	617	2	AAW00636 N-termina
59	3154	52.9	588	2	AAW00635 N-termina
60	3138	52.6	588	2	AAW00644 N-termina
61	3125	52.4	622	2	AAW25463 Human CRT
62	3038	50.9	591	2	AAW97384 A catalyt
63	2581	43.3	500	6	AAO29840 Human tel
64	2573	43.1	499	6	ABB99678 Amino aci
65	2253	37.8	1131	7	ADD21415 Frog TERT
66	2178	36.5	436	6	ABB99680 Splice va
67	2178	36.5	438	2	AAW25462 Human CRT
68	2167.5	36.3	743	8	ADG90607 Dog TERT
69	2167	36.3	463	6	ABB99679 Splice va
70	2163.5	36.3	437	2	AAW25461 Human CRT
71	2079	34.9	564	2	AAW56109 Human tel
72	2007	33.7	575	8	ADG90605 Rat TERT
73	1863	31.2	348	2	AAW00645 Truncated
74	1541	25.8	538	2	AAW47001 Glutathio
75	1506	25.3	514	2	AAW47004 Glutathio
76	1484	24.9	291	6	AAO29774 hTERT MHC
77	1447	24.3	531	2	AAW47002 Glutathio
78	1438	24.1	283	2	AAW43128 Human tel
79	1417.5	23.8	516	2	AAW47005 Glutathio
80	1338.5	22.4	514	2	AAW47003 Glutathio
81	1293	21.7	250	8	ADG85224 Human tel
82	1090	18.3	259	2	AAW46998 Human tel
83	902	15.1	174	6	AAO29775 hTERT MHC
84	708	11.9	379	4	AAE00431 Consensus
85	667	11.2	174	6	ABB99681 Splice va
86	645	10.8	131	2	AAW97385 Amino aci
87	588	9.5	988	2	AAW56107 S. pombe
88	565	9.5	816	8	ADG70135 HIV RT/HT
89	555	9.3	803	8	ADG70131 HIV RT/HT
90	555	9.3	816	8	ADG70133 HIV RT/HT
91	555	9.3	108	6	ABB99682 Splice va
92	535	9.0	100	5	ABG71628 hTERT fra
93	534	9.0	100	5	ABG71627 hTERT fra
94	503	8.4	100	5	ABG70112 HIV-1 RT/
95	478	8.0	576	8	ADG70120 HIV RT/HT
96	478	8.0	592	8	ADG70121 HIV RT/HT
97	478	8.0	605	8	ADG70122 HIV RT/HT
98	461	7.7	816	8	ADG70132 HIV RT/HT

99 459 7.7 456 8 ADG70126
100 459 7.7 462 8 ADG70124

Adg70126 HIV RT/ht
Adg70124 HIV RT/ht

ALIGNMENTS

RESULT 1

AAW56113

ID AAW56113 standard; protein; 1132 AA.

XX AC

XX AAW56113;

XX 13-AUG-1998 (first entry)

XX DT

XX DE

XX Human telomerase reverse transcriptase protein refined sequence.

XX KW

XX Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis; prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.

XX OS

XX Homo sapiens.

XX FN

XX GB2317891-A.

XX XX

XX 08-APR-1998.

XX PD

XX 01-OCT-1997;

XX 97GB-00020890.

XX PR

XX 01-OCT-1996;

XX 96US-00724643.

XX PR

XX 18-APR-1997;

XX 97US-00844419.

XX PR

XX 25-APR-1997;

XX 97US-00846017.

XX PR

XX 06-MAY-1997;

XX 97US-00851843.

XX PR

XX 09-MAY-1997;

XX 97US-00854050.

XX PR

XX 14-AUG-1997;

XX 97US-00911312.

XX PR

XX 14-AUG-1997;

XX 97US-00912951.

XX PR

XX 14-AUG-1997;

XX 97US-00915503.

XX XX

XX (GERO-) GERON CORP.

XX (UYTE-) UNIV TECHNOLOGY CORP.

XX PA

XX Cech TR, Lingner J, Nakamura T, Chapman KB, Morin GB, Harley CB;

XX PI Andrews WH;

XX FI

XX WPI: 1998-171633/16.

XX DR

XX N-PSDB; AAW22428.

XX PT

XX Pure and recombinant human Telomerase Reverse Transcriptase and its

XX variants - are useful in the diagnosis, prognosis and treatment of cell

XX proliferation conditions especially cancer and ageing.

XX PS

XX Example 1; Fig 74; 387pp; English.

XX CC

XX The present sequence represents human telomerase reverse transcriptase

XX (hTERT), which is a ribonucleoprotein. The present invention also

XX describes the following methods: (A) determining whether a test compound

XX is a modulator of hTERT by detecting the change in hTERT recombinant

XX protein or polynucleotide, on administration of the compound; (B)

XX preparation of recombinant telomerase by contacting a protein preparation

XX of hTERT with a telomerase RNA component; (C) detection of the hTERT RNA or

XX protein in a sample by binding a relevant probe to the sample and

XX detecting the complex formed or in the case of RNA detection, amplifying

XX the product and correlating the presence of complex or amplification

XX product with presence of hTERT in the sample; and (D) increasing the

XX proliferation of a vertebrate cell by increasing hTERT expression; and (E)

XX the use of an agent that causes an increase in cell vertebrate cell

XX proliferation to create a medicament that inhibits ageing. A protein

XX preparation of hTERT and the polynucleotide encoding hTERT can be used in

XX the manufacture of medicaments for inhibiting the effect of ageing or

XX cancer. Inhibitors of telomerase activity can be used to treat conditions

XX that are associated with high telomerase activity. A protein preparation

XX of hTERT can also be used in the new methods

XX Sequence 1132 AA;

XX SQ

Query Match 100.0%; Score 5963; DB 2; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPRAPRCRAVRSLLRSHYREVLPATFVRLGPGWRLVORGDPAAFRALVAQCVCVWP 60
Db 1 MPRAPRCRAVRSLLRSHYREVLPATFVRLGPGWRLVORGDPAAFRALVAQCVCVWP 60
Qy 61 DARPPPAAPSFRQVSCUKELVARVLQRLCERGAKNVLAFGFALDGDGARGGPEAFTTSVR 120
Db 61 DARPPPAAPSFRQVSCUKELVARVLQRLCERGAKNVLAFGFALDGDGARGGPEAFTTSVR 120
Qy 121 SYLPNTVTDALRGSGAMGLLLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Db 121 SYLPNTVTDALRGSGAMGLLLRRVGGDDVLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180
Qy 181 ATQARPPPHASGPRRRRLGCRAMNHSVREAGVPLGLPAPGARRRGGSASRSLPLPKRPRR 240
Db 181 ATQARPPPHASGPRRRRLGCRAMNHSVREAGVPLGLPAPGARRRGGSASRSLPLPKRPRR 240
Qy 241 GAAPEPERTVGGGSAHPGRTGRGSDRGFCVSPARPAEATSLGALSSTRHSHSVG 300
Db 241 GAAPEPERTVGGGSAHPGRTGRGSDRGFCVSPARPAEATSLGALSSTRHSHSVG 300
Qy 301 RQHHAGPSTSRPPRPMWDTPCPPVYAEAKHFLYSSGDKQLRPSFLSSLRPSLTGARRL 360
Db 301 RQHHAGPSTSRPPRPMWDTPCPPVYAEAKHFLYSSGDKQLRPSFLSSLRPSLTGARRL 360
Qy 361 VETIFLGSRPMPGTPRRRLPRLPORYWQMRPLFLELGNHAQCYPGYVLLKTHCPLRAAVT 420
Db 361 VETIFLGSRPMPGTPRRRLPRLPORYWQMRPLFLELGNHAQCYPGYVLLKTHCPLRAAVT 420
Qy 421 PAAGVCAREKPOGSAVAPEEEDTPRELVLQRLHSSPMQVYGVFVRACLRLVPPGLWGS 480
Db 421 PAAGVCAREKPOGSAVAPEEEDTPRELVLQRLHSSPMQVYGVFVRACLRLVPPGLWGS 480
Qy 481 RHNERRLRNTKTFISLGKHAQLSLOELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 540
Db 481 RHNERRLRNTKTFISLGKHAQLSLOELTWKMSVRDCAWLRRSPGVGCVPAAEHRLREEI 540
Qy 541 LAKFLHLMSSVYVVELLRSFFYVTTTFOKNRLFFYRPSVMSKLOQSIGIRHKLKRVQURE 600
Db 541 LAKFLHLMSSVYVVELLRSFFYVTTTFOKNRLFFYRPSVMSKLOQSIGIRHKLKRVQURE 600
Qy 601 LSEAEVQHEAREPALLTSRLRFTPKDGLRPIVNMNDYVVGARTFRREKRAERLTSRVKA 660
Db 601 LSEAEVQHEAREPALLTSRLRFTPKDGLRPIVNMNDYVVGARTFRREKRAERLTSRVKA 660
Qy 661 LFSVLNYERARRPGLLGASVLGLDDIHRWRTFVLVRVRAQDPPPELVFVKVDVTGAYDTI 720
Db 661 LFSVLNYERARRPGLLGASVLGLDDIHRWRTFVLVRVRAQDPPPELVFVKVDVTGAYDTI 720
Qy 721 PQDLTEVIASIIKPONTYCVRRYAVVQKAAHGVKAFKSHVSTLTLDQYMQFVAHL 780
Db 721 PQDLTEVIASIIKPONTYCVRRYAVVQKAAHGVKAFKSHVSTLTLDQYMQFVAHL 780
Qy 781 QETSPLDADVIEOSSINEASSGLFDVFLRFMCHHVRIRKGSVVOCGIPQGSILSTL 840
Db 781 QETSPLDADVIEOSSINEASSGLFDVFLRFMCHHVRIRKGSVVOCGIPQGSILSTL 840
Qy 841 LCSLCYGDMMENKLFAGIRRRDGLLLRLVDDFLLVTPHLLTHAKTFLRLVRGVPYGCVVNL 900
Db 841 LCSLCYGDMMENKLFAGIRRRDGLLLRLVDDFLLVTPHLLTHAKTFLRLVRGVPYGCVVNL 900
Qy 901 RKTWNVPFVEDEALGGTAFVQMPAHGLFPMCGLLDTRTLEVSQSDYSYARTSIRASVTF 960
Db 901 RKTWNVPFVEDEALGGTAFVQMPAHGLFPMCGLLDTRTLEVSQSDYSYARTSIRASVTF 960
Qy 961 NRGFAGRNRRKLFGLVRLKCHSLFLDLQVNSLOTVCTNLYKILLQAYRHFACVLOLP 1020
Db 961 NRGFAGRNRRKLFGLVRLKCHSLFLDLQVNSLOTVCTNLYKILLQAYRHFACVLOLP 1020

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:28:46 ; Search time 5471 Seconds
(without alignments)
7875.846 Million cell updates/sec

Title: US-08-974-584C-118
Perfect score: 5963
Sequence: 1 MPRAPCRRAVRSLLRSHRYE.....TALEAANPALPSPDKTILD 1132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US08974584/runat_25072005.103113.9073/app_query.fasta_1.1287
-DB=EST -QFMT=fastap -SUFFIX=n2p.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US08974584 @CGN 1 1 3202 @runat_25072005.103113.9073 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gse1:
9: gb_gse2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2636	44.2	1826	9 AY407349	AY407349 Homo sapi
2	2159	36.2	1584	9 AY407350	AY407350 Pan trogl
3	1796	30.1	1835	9 AY407351	AY407351 Mus muscu
4	1403	23.5	925	4 BM453198	BM453198 AGENCOURT
5	1025	17.2	851	5 BU702370	BU702370 UI-M-F10-
6	872	14.6	851	4 BG917907	BG917907 602820830
7	773	13.0	492	4 BM824748	BM824748 K-EST0096
8	765.5	12.8	664	5 BQ258274	BQ258274 NISC kp11
9	679	11.4	1424	3 CR688161	CR688161 Tetraodon

10	667.5	11.2	649	7	CF531069	CF531069
11	666.5	11.2	688	7	CF531121	CF531121
12	640	10.7	389	1	AA281296	AA281296
13	610	10.2	599	7	BB618671	BB618671
14	567	9.5	409	7	CN274427	CN274427
15	566	9.5	866	2	BE371943	BE371943
16	554	9.3	614	2	BB651920	BB651920
17	515	8.6	753	5	BU452535	BU452535
18	514	8.6	775	4	BI388013	BI388013
19	510.5	8.6	1872	9	CL979127	CL979127
20	481.5	8.0	846	7	CN505902	CN505902
21	478.5	8.0	880	5	BU377259	BU377259
22	462	7.7	668	6	CA380121	CA380121
23	456	7.6	724	5	EX889962	EX889962
24	446.5	7.5	632	6	CA353864	CA353864
25	437.5	7.3	715	2	BE396925	BE396925
26	434	7.3	703	5	EX886589	EX886589
27	433	7.3	344	7	CF531258	CF531258
28	425	7.1	409	1	AA311750	AA311750
29	424	7.1	835	5	BU111946	BU111946
30	419	7.0	679	2	BE396606	BE396606
31	416	7.0	649	2	BE514070	BE514070
32	405	6.8	610	2	BE514188	BE514188
33	395.5	6.6	696	5	BU139751	BU139751
34	387	6.5	343	6	BY783093	BY783093
35	384	6.4	338	6	BY784804	BY784804
36	380	6.4	336	6	BY775178	BY775178
37	377	6.3	347	2	AW244516	AW244516
38	374	6.3	326	5	BY149368	BY149368
39	373	6.3	641	8	AZ972318	AZ972318
40	366	6.1	619	7	CK392784	CK392784
41	351.5	5.9	779	2	BE268183	BE268183
42	348.5	5.8	875	5	BU122597	BU122597
43	348.5	5.8	930	7	CO014076	CO014076
44	311	5.2	983	7	CO028055	CO028055
45	310.5	5.2	1197	9	CL948838	CL948838
46	304	5.1	1023	7	CO024489	CO024489
47	291.5	4.9	774	7	CF547484	CF547484
48	291	4.9	687	6	CD310596	CD310596
49	284.5	4.8	813	5	BU224024	BU224024
50	283	4.7	646	5	EX882610	EX882610
51	283	4.7	731	5	EX315053	EX315053
52	274	4.6	774	3	CR717591	CR717591
53	273.5	4.6	651	5	BM363763	BM363763
54	265	4.4	570	4	BM521744	BM521744
55	265	4.4	1455	8	CC190951	CC190951
56	264	4.4	534	8	AQ397020	AQ397020
57	262	4.4	1554	8	CC190875	CC190875
58	261	4.4	1230	8	CC211572	CC211572
59	258.5	4.3	502	2	BF251764	BF251764
60	250	4.2	1020	9	CNS05FIM	CNS05FIM
61	249.5	4.2	851	9	CG147626	CG147626
62	235	3.9	724	5	EX315052	EX315052
63	228.5	3.8	534	2	AW318894	AW318894
64	228.5	3.8	568	5	EX521269	EX521269
65	221.5	3.7	393	2	BF251775	BF251775
66	221	3.7	386	5	EX088059	EX088059
67	217.5	3.6	1009	9	CNS0730U	CNS0730U
68	212	3.6	696	8	BM962910	BM962910
69	208	3.5	1412	4	BM555089	BM555089
70	207.5	3.5	1913	9	CL978828	CL978828
71	206	3.5	954	9	AG030799	AG030799
72	205.5	3.4	625	8	B27802	B27802
73	205	3.4	613	2	BF597086	BF597086
74	204	3.4	813	4	EG198331	EG198331
75	204	3.4	839	7	CO814083	CO814083
76	201	3.4	817	7	CO014075	CO014075
77	199	3.3	2259	9	AG280420	AG280420
78	196	3.3	1146	5	BF082688	BF082688
79	194	3.3	1146	5	BQ683700	BQ683700
80	193.5	3.2	2868	9	CL982351	CL982351
81	193	3.2	3953	3	AK046822	AK046822
82	192	3.2	1329	9	CL515191	CL515191

83	191	3.2	366	2	BF511837	BF511837 UI-H-B14-
84	191	3.2	1732	9	CL489127	CL489127 SAIL 519
85	191	3.2	2295	9	AG280482	AG280482 Mus muscu
86	190	3.2	916	9	CG247145	CG247145 OG3BAL7TH
c 87	190	3.2	1370	5	BQ651339	BQ651339 AGENCOURT
c 88	190	3.2	8004	9	CU968201	CU968201 OBIFFC016
89	189	3.2	2020	9	AG304986	AG304986 Mus muscu
90	189	3.2	4323	9	CL982276	CL982276 OeIFSC047
91	188.5	3.2	1774	4	BG743567	BG743567 602633914
c 92	187.5	3.1	1145	9	AG070955	AG070955 Pan trogl
c 93	187.5	3.1	1765	9	AY416688	AY416688 Homo sapi
94	187	3.1	997	5	BQ935457	BQ935457 AGENCOURT
95	186.5	3.1	1781	3	CR621614	CR621614 full-leng
96	186.5	3.1	1796	3	CR593631	CR593631 full-leng
97	186.5	3.1	1798	3	CR591255	CR591255 full-leng
98	186.5	3.1	1804	3	CR615050	CR615050 full-leng
c 99	186.5	3.1	2181	3	AK032335	AK032335 Mus muscu
c 100	186	3.1	1212	8	BZ556479	BZ556479 pacs1-60_

RESULT 1	AY407349	1826 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY407349				
DEFINITION	Homo sapiens TERT gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
ACCESSION	AY407349				
VERSION	AY407349.1	GI:39763320			
KEYWORDS	GSS.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science	302	(5652)	1960-1963	(2003)
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1826)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				
JOURNAL	Submitted	(16-NOV-2003)	Celera Genomics	45 West Gude Drive,	
COMMENT	Rockville, MD 20850, USA				
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
source	Location/Qualifiers				
gene	1..1826				
ORIGIN	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	<1..>1826				
	/gene="TERT"				
	/locus_tag="HCM2861"				

Alignment Scores:					
Pred. No.:	2,91e-172	Length:	1826		
Score:	2636.00	Matches:	524		
Percent Similarity:	86.43%	Conservative:	1		
Best Local Similarity:	86.33%	Mismatches:	82		
Query Match:	44.21%	Indels:	0		
DB:	9	Gaps:	0		

US-08-974-584C-118 (1-1132) x AY407349 (1-1826)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 14:12:42 ; Search time 1052 Seconds
(without alignments)
6958.011 Million cell updates/sec

Title: US-08-974-584C-118
Perfect score: 5963
Sequence: 1 MFAPRCRAVRSLLRSHYRE.....TALEAANPALPSPDKTILD 1132

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7277826 seqs, 3231319505 residues
Total number of hits satisfying chosen parameters: 14555652

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US08974584/runat_25072005_103115_9161/app_query.fasta_1.1287
-DB=Published_Applications_NA -GFM=fastap -SUFFIX=n2p.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US08974584 @CGN 1.1 487 @runat_25072005_103115_9161
-NCPUI=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEVTIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5963	100.0	3453	14	US-10-205-629-1	Sequence 1, Appli
2	5963	100.0	4037	17	US-10-325-810-343	Sequence 343, App
3	5963	100.0	4037	20	US-10-877-124-343	Sequence 343, App
4	5963	100.0	4037	20	US-10-877-022-343	Sequence 343, App
5	5963	100.0	4037	21	US-10-877-146-343	Sequence 343, App
6	5963	100.0	13766	16	US-10-105-616-1	Sequence 1, Appli
7	5952	99.8	3396	9	US-09-749-7288-32	Sequence 32, Appli
8	5952	99.8	3396	20	US-10-877-124-638	Sequence 638, App
9	5952	99.8	3396	20	US-10-877-124-639	Sequence 639, App
10	5952	99.8	3396	20	US-10-877-124-640	Sequence 640, App
11	5952	99.8	3396	20	US-10-877-124-641	Sequence 641, App
12	5952	99.8	3396	20	US-10-877-124-642	Sequence 642, App
13	5952	99.8	3396	20	US-10-877-022-638	Sequence 638, App
14	5952	99.8	3396	20	US-10-877-022-639	Sequence 639, App
15	5952	99.8	3396	20	US-10-877-022-640	Sequence 640, App
16	5952	99.8	3396	20	US-10-877-022-641	Sequence 641, App
17	5952	99.8	3396	20	US-10-877-022-642	Sequence 642, App
18	5952	99.8	3396	21	US-10-877-146-638	Sequence 638, App
19	5952	99.8	3396	21	US-10-877-146-639	Sequence 639, App
20	5952	99.8	3396	21	US-10-877-146-640	Sequence 640, App
21	5952	99.8	3396	21	US-10-877-146-641	Sequence 641, App
22	5952	99.8	3396	21	US-10-877-146-642	Sequence 642, App
23	5952	99.8	3399	19	US-10-384-339C-28	Sequence 28, Appli
24	5952	99.8	3451	20	US-10-877-124-721	Sequence 721, App
25	5952	99.8	3451	20	US-10-877-022-721	Sequence 721, App
26	5952	99.8	3451	21	US-10-877-146-721	Sequence 721, App
27	5952	99.8	4015	9	US-09-733-294A-3	Sequence 3, Appli
28	5952	99.8	4015	9	US-09-990-080-1	Sequence 1, Appli
29	5952	99.8	4015	9	US-09-843-676-224	Sequence 224, App
30	5952	99.8	4015	9	US-09-953-052-1	Sequence 1, Appli
31	5952	99.8	4015	14	US-10-053-758-224	Sequence 224, App
32	5952	99.8	4015	14	US-10-208-243-1	Sequence 1, Appli
33	5952	99.8	4015	14	US-10-054-295-224	Sequence 224, App
34	5952	99.8	4015	14	US-10-054-611-224	Sequence 224, App
35	5952	99.8	4015	14	US-10-054-663-1	Sequence 1, Appli
36	5952	99.8	4015	14	US-10-044-692-1	Sequence 1, Appli
37	5952	99.8	4015	15	US-10-044-539-1	Sequence 1, Appli
38	5952	99.8	4015	17	US-10-325-810-1	Sequence 1, Appli
39	5952	99.8	4015	17	US-10-388-578-1	Sequence 1, Appli
40	5952	99.8	4015	17	US-10-439-565-1	Sequence 1, Appli
41	5952	99.8	4015	19	US-10-602-441-1	Sequence 1, Appli
42	5952	99.8	4015	19	US-10-389-431-1	Sequence 1, Appli
43	5952	99.8	4015	20	US-10-877-124-1	Sequence 1, Appli
44	5952	99.8	4015	20	US-10-877-022-1	Sequence 1, Appli
45	5952	99.8	4015	21	US-10-831-266-3	Sequence 3, Appli
46	5952	99.8	4015	21	US-10-432-034-1	Sequence 1, Appli
47	5952	99.8	4015	21	US-10-831-267-3	Sequence 3, Appli
48	5952	99.8	4015	21	US-10-877-146-1	Sequence 1, Appli
49	5952	99.8	4015	21	US-10-917-853-1	Sequence 1, Appli
50	5952	99.8	4015	21	US-10-918-739-1	Sequence 1, Appli
51	5952	99.8	4015	21	US-10-637-443-1	Sequence 1, Appli
52	5952	99.8	4015	21	US-10-794-514A-4	Sequence 4, Appli
53	5952	99.8	4015	22	US-10-143-536-1	Sequence 1, Appli
54	5952	99.8	4027	21	US-10-840-455-2	Sequence 2, Appli
55	5945	99.7	4027	17	US-10-385-882-1	Sequence 1, Appli
56	5936	99.5	8742	16	US-10-105-616-6	Sequence 6, Appli
57	5769.5	96.8	3369	21	US-10-794-514A-339	Sequence 339, App
58	5584.5	93.7	3855	14	US-10-044-692-4	Sequence 4, Appli
59	5584.5	93.7	3855	15	US-10-044-539-4	Sequence 4, Appli
60	5584.5	93.7	3855	17	US-10-325-810-4	Sequence 4, Appli
61	5584.5	93.7	3855	20	US-10-877-124-4	Sequence 4, Appli
62	5584.5	93.7	3855	20	US-10-877-022-4	Sequence 4, Appli
63	5584.5	93.7	3855	21	US-10-877-146-4	Sequence 4, Appli
64	5570	93.4	4029	9	US-09-843-676-173	Sequence 173, App
65	5570	93.4	4029	10	US-09-438-486-173	Sequence 173, App

C

```
66 5570 93.4 4029 14 US-10-053-758-173 Sequence 173, App
67 5570 93.4 4029 14 US-10-054-295-173 Sequence 173, App
68 5570 93.4 4029 14 US-10-054-611-173 Sequence 173, App
69 5570 93.4 4029 17 US-10-325-810-292 Sequence 292, App
70 5570 93.4 4029 20 US-10-877-124-292 Sequence 292, App
71 5570 93.4 4029 21 US-10-877-022-292 Sequence 292, App
72 5570 93.4 4029 21 US-10-877-146-292 Sequence 292, App
73 5398.5 90.5 3255 21 US-10-794-514A-337 Sequence 337, App
74 3619 60.7 4170 19 US-10-602-441-5 Sequence 5, Appli
75 3496 58.6 3426 13 US-10-602-441-3 Sequence 3, Appli
76 3496 58.6 3426 21 US-10-794-514A-6 Sequence 6, Appli
77 3496 58.6 3496 20 US-10-862-698-1 Sequence 11, Appl
78 3125 52.4 1866 14 US-10-294-778-11 Sequence 43, Appl
79 2789 46.8 26414 21 US-09-733-294A-30 Sequence 30, Appl
80 2789 46.8 51552 9 US-09-733-294A-30 Sequence 6, Appli
81 2782 46.7 4200 14 US-10-044-692-6 Sequence 6, Appli
82 2782 46.7 4200 15 US-10-044-539-6 Sequence 6, Appli
83 2782 46.7 4321 17 US-10-325-810-6 Sequence 6, Appli
84 2782 46.7 4335 20 US-10-877-124-6 Sequence 6, Appli
85 2782 46.7 4335 20 US-10-877-022-6 Sequence 6, Appli
86 2782 46.7 4335 21 US-10-877-146-6 Sequence 6, Appli
87 2782 46.7 15418 9 US-09-783-203-1 Sequence 1, Appli
88 2782 46.7 15418 9 US-09-994-427A-1 Sequence 1, Appli
89 2782 46.7 15418 10 US-09-995-419A-1 Sequence 1, Appli
90 2782 46.7 15418 14 US-10-141-220-1 Sequence 1, Appli
91 2782 46.7 15418 14 US-10-123-969-1 Sequence 1, Appli
92 2782 46.7 15418 14 US-10-206-447-1 Sequence 1, Appli
93 2782 46.7 15418 18 US-10-674-836-1 Sequence 1, Appli
94 2782 46.7 15418 19 US-10-811-012-1 Sequence 1, Appli
95 2661.5 44.6 2176 14 US-10-044-692-3 Sequence 3, Appli
96 2661.5 44.6 2176 15 US-10-044-539-3 Sequence 3, Appli
97 2661.5 44.6 2176 17 US-10-325-810-3 Sequence 3, Appli
98 2661.5 44.6 2176 20 US-10-877-124-3 Sequence 3, Appli
99 2661.5 44.6 2176 20 US-10-877-022-3 Sequence 3, Appli
100 2661.5 44.6 2176 21 US-10-877-146-3 Sequence 3, Appli
```

ALIGNMENTS

RESULT 1

```
US-10-205-629-1
; Sequence 1, Application US/10205629
; Publication No. US2003049236A1
; GENERAL INFORMATION:
; APPLICANT: Kassem, Moustapha
; APPLICANT: Jensen, Thomas
; APPLICANT: Rattan, Suresh
; TITLE OF INVENTION: Immortalized Stem Cells
; FILE REFERENCE: 006148.00002
; CURRENT APPLICATION NUMBER: US/10/205,629
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 60/315939
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: PA 2001 01148
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-629-1
```

Alignment Scores:

```
Pred. No.: 0 Length: 3453
Score: 5963.00 Matches: 1132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
```

US-08-974-584C-118 (1-1132) x US-10-205-629-1 (1-3453)

```
Qy 1 MetProArgAlaProArgCysArgAlaValArgSerLeuLeuArgSerHisTyrArgGlu 20
Db 10 ATGCGCGCGCTCCCGCTGCCGAGCGTGCGCTCCCTGCTGCGCAGCACTACTACCGCAG 69
Qy 21 ValLeuProLeuAlaThrPheValArgArgLeuGlyProGlnGlyTyrArgLeuValGln 40
Db 70 GTGTCGCGCTGGCGACGTTCTGTCGCGCGCTGGGGGCCCCAGAGGCTGGCGCTGGTGAG 129
Qy 41 ArgGlyAspProAlaAlaPheArgAlaLeuValAlaGlnCysValLeuValCysValProTrp 60
Db 130 CGCGGGGACCGCGCGGCTTCCCGCGCTGCTGGCGAGTGCCTGGGTGGTGGTGGTGGTGG 189
Qy 61 AspAlaArgProProProAlaAlaProSerPheArgGlnValSerCysLeuLysGluLeu 80
Db 190 GAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 249
Qy 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaLysAsnValLeuAlaPheGly 100
Db 250 GTGCGCGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 309
Qy 101 PheAlaLeuLeuAspGlyAlaArgGlyGlyProProGluAlaPheThrThrSerValArg 120
Db 310 TTCGCGCTGTGGACGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCGGGGGCGCGGGGG 369
Qy 121 SerTyrLeuProAsnThrValThrAspAlaLeuArgGlySerGlyAlaTyrGlyLeuLeu 140
Db 370 AGCTACCTGCGCCACACGCTGACCGACGACTGCGGGGGAGCGGGGGCGGGGGCGGGGG 429
Qy 141 LeuArgArgValGlyAspAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
Db 430 CTGCGCGCGTGGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
Qy 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
Db 490 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 549
Qy 181 AlaThrGlnAlaArgProProHisAlaSerGlyProArgArgArgLeuGlyCysGlu 200
Db 550 GCCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 609
Qy 201 ArgAlaTrpAsnHisSerValArgGluAlaGlyValProLeuGlyLeuProAlaProGly 220
Db 610 CGGGCTTGGAAACCATAGCGTCAGGAGCGCGGGGTCCCTGCGGCTGCCAGAGCCCGGG 669
Qy 221 AlaArgArgArgGlyGlySerAlaSerArgSerLeuProLeuProLysArgProArgArg 240
Db 670 GCGAGGAGCGCGGGGCGAGTGCAGCGAGTCTGCGGTGCGGCGCGCGCGCGCGCGCGCG 729
Qy 241 GlyAlaAlaProGluProGluArgThrProValGlyGlnGlySerTrpAlaHisProGly 260
Db 730 GCGCTGCGCTGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 789
Qy 261 ArgThrArgGlyProSerAspArgGlyPheCysValValSerProAlaArgProAlaGlu 280
Db 790 AGGACGCGTGGACCGAGTACCGTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
Qy 281 GluAlaThrSerLeuGluGlyAlaLeuSerGlyThrArgHisSerHisProSerValGly 300
Db 850 GAAGCCACCTCTTTGGAGGGTGGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCG 909
Qy 301 ArgGlnHisAlaGlyProProSerThrSerArgProProArgProTrpAspThrPro 320
Db 910 CGCCAGCACCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 969
Qy 321 CysProProValTyrAlaGluThrLysHisPheLeuTyrSerSerGlyAspLysGluGln 340
Db 970 TGTCCCGCGGTGTACGCGGAGACCAAGCACTTCTCTACTCTCTCAGGCGCAAGAGGAG 1029
Qy 341 LeuArgProSerPheLeuLeuSerSerLeuArgProSerLeuThrGlyValArgArgLeu 360
Db 1030 CTGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
Qy 361 ValGluThrIlePheLeuGlySerArgProTrpMetProGlyThrProArgArgLeuPro 380
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:31:01 ; Search time 309 Seconds
(without alignments)
5994.386 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRLSLRSHYRE.....TALEAANPALPDKFTILD 1132

Scoring table: *BLOSUM62

Qa/cgn2_1/USPTO spool/US08974584/runat_25072005_103113_9083/app query.fasta_1.1287
-Qa/cgn2_1/USPTO spool/US08974584/runat_25072005_103113_9083/app query.fasta_1.1287
-DB=issued_Patents_NA -Qfmt=fastep -SUFFIX=n2p.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US08974584 @CGN 1.1 59 @runat_25072005_103113_9083 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Qa/cgn2_1/USPTO spool/US08974584/runat_25072005_103113_9083/app query.fasta_1.1287
-DB=issued_Patents_NA -Qfmt=fastep -SUFFIX=n2p.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US08974584 @CGN 1.1 59 @runat_25072005_103113_9083 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:
6: /cgn2_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5963	100.0	4037	3	US-08-974-549A-343
2	5963	100.0	4037	3	US-09-402-181B-343
3	5963	100.0	4037	4	US-09-721-456-343
4	5952	99.8	3396	3	US-08-974-549A-638
5	5952	99.8	3396	3	US-08-974-549A-639
6	5952	99.8	3396	3	US-08-974-549A-640
7	5952	99.8	3396	3	US-08-974-549A-641
8	5952	99.8	3396	3	US-08-974-549A-642
9	5952	99.8	3396	4	US-09-721-456-638
10	5952	99.8	3396	4	US-09-721-456-639
11	5952	99.8	3396	4	US-09-721-456-640
12	5952	99.8	3396	4	US-09-721-456-641

13	5952	99.8	3396	4	US-09-721-456-642	Sequence 642, App
14	5952	99.8	3451	3	US-08-974-549A-721	Sequence 721, App
15	5952	99.8	3451	3	US-09-721-456-721	Sequence 721, App
16	5952	99.8	4015	3	US-08-851-843A-224	Sequence 224, App
17	5952	99.8	4015	3	US-08-974-549A-1	Sequence 1, Appli
18	5952	99.8	4015	3	US-08-854-050-224	Sequence 224, App
19	5952	99.8	4015	3	US-09-430-323-224	Sequence 224, App
20	5952	99.8	4015	3	US-09-572-423B-3	Sequence 3, Appli
21	5952	99.8	4015	3	US-09-128-354-1	Sequence 1, Appli
22	5952	99.8	4015	3	US-09-675-321-1	Sequence 1, Appli
23	5952	99.8	4015	3	US-09-052-919-1	Sequence 1, Appli
24	5952	99.8	4015	3	US-08-912-951-1	Sequence 1, Appli
25	5952	99.8	4015	4	US-09-733-294A-3	Sequence 3, Appli
26	5952	99.8	4015	4	US-09-402-181B-1	Sequence 1, Appli
27	5952	99.8	4015	4	US-09-721-456-1	Sequence 1, Appli
28	5952	99.8	4015	4	US-09-953-052-1	Sequence 1, Appli
29	5952	99.8	4015	4	US-09-465-491-1	Sequence 1, Appli
30	5952	99.8	4015	4	US-09-949-016-455	Sequence 455, App
31	5952	99.8	4015	4	US-09-601-645A-10	Sequence 10, Appl
32	5934	99.5	4016	4	US-09-949-016-5841	Sequence 5841, Ap
33	5584.5	93.7	3855	3	US-08-974-549A-4	Sequence 4, Appli
34	5584.5	93.7	3855	4	US-08-912-951-4	Sequence 4, Appli
35	5584.5	93.7	3855	4	US-09-402-181B-4	Sequence 4, Appli
36	5584.5	93.7	3855	4	US-09-721-456-4	Sequence 4, Appli
37	5570	93.4	4029	3	US-08-851-843A-173	Sequence 173, App
38	5570	93.4	4029	3	US-08-974-549A-292	Sequence 292, App
39	5570	93.4	4029	3	US-08-854-050-173	Sequence 173, App
40	5570	93.4	4029	3	US-09-430-323-173	Sequence 173, App
41	5570	93.4	4029	4	US-09-402-181B-292	Sequence 292, App
42	5570	93.4	4029	4	US-09-721-456-292	Sequence 292, App
43	3496	58.6	3496	4	US-09-042-460-1	Sequence 1, Appli
44	3125	52.4	1866	4	US-09-582-924B-11	Sequence 11, Appli
45	2789	46.8	44952	4	US-09-949-016-12197	Sequence 12197, A
46	2789	46.8	44960	4	US-09-949-016-17583	Sequence 17583, A
47	2789	46.8	51552	4	US-09-733-294A-30	Sequence 30, Appli
48	2782	46.7	4200	4	US-08-912-951-6	Sequence 6, Appli
49	2782	46.7	4321	4	US-09-402-181B-6	Sequence 6, Appli
50	2782	46.7	4335	3	US-08-974-549A-6	Sequence 6, Appli
51	2782	46.7	4335	3	US-09-721-456-6	Sequence 6, Appli
52	2782	46.7	15418	4	US-09-783-203-1	Sequence 1, Appli
53	2782	46.7	15418	4	US-09-994-427A-1	Sequence 1, Appli
54	2782	46.7	15418	4	US-09-244-438-1	Sequence 1, Appli
55	2661.5	44.6	2176	3	US-08-974-549A-3	Sequence 3, Appli
56	2661.5	44.6	2176	4	US-08-912-951-3	Sequence 3, Appli
57	2661.5	44.6	2176	4	US-09-402-181B-3	Sequence 3, Appli
58	2661.5	44.6	2176	4	US-09-721-456-3	Sequence 3, Appli
59	2582.5	43.3	2171	3	US-08-851-843A-100	Sequence 100, App
60	2582.5	43.3	2171	3	US-08-974-549A-266	Sequence 266, App
61	2582.5	43.3	2171	3	US-08-854-050-100	Sequence 100, App
62	2582.5	43.3	2171	3	US-09-430-323-100	Sequence 100, App
63	2582.5	43.3	2171	4	US-09-402-181B-266	Sequence 266, App
64	2582.5	43.3	2171	4	US-09-721-456-266	Sequence 266, App
65	2582.5	43.3	2171	4	US-09-766-253-100	Sequence 100, App
66	2178	36.5	1314	4	US-09-582-924B-9	Sequence 9, Appli
67	2163.5	32.7	1311	4	US-09-582-924B-1	Sequence 1, Appli
68	1356.5	22.7	7498	4	US-09-244-438-2	Sequence 2, Appli
69	1092	18.3	601	4	US-09-949-016-26770	Sequence 26770, A
70	808	13.6	601	4	US-09-949-016-26771	Sequence 26771, A
71	640	10.7	389	3	US-08-851-843A-62	Sequence 62, Appli
72	640	10.7	389	3	US-08-974-549A-8	Sequence 8, Appli
73	640	10.7	389	3	US-08-854-050-62	Sequence 62, Appli
74	640	10.7	389	3	US-09-430-323-62	Sequence 62, Appli
75	640	10.7	389	4	US-08-912-951-8	Sequence 8, Appli
76	640	10.7	389	4	US-09-402-181B-8	Sequence 8, Appli
77	640	10.7	389	4	US-09-721-456-8	Sequence 8, Appli
78	640	10.7	389	4	US-09-766-253-62	Sequence 62, Appli
79	453	7.6	2651	4	US-09-042-460-5	Sequence 5, Appli
80	425	7.1	409	4	US-09-733-294A-31	Sequence 31, Appli
81	418.5	7.0	3279	3	US-08-851-843A-1	Sequence 1, Appli
82	418.5	7.0	3279	3	US-08-974-549A-109	Sequence 109, App
83	418.5	7.0	3279	3	US-08-854-050-1	Sequence 1, Appli
84	418.5	7.0	3279	3	US-09-430-323-1	Sequence 1, Appli
85	418.5	7.0	3279	4	US-08-912-951-109	Sequence 109, App

86 418.5 7.0 3279 4 US-09-402-181B-109
Sequence 109, App
87 418.5 7.0 3279 4 US-09-721-456-109
Sequence 109, App
88 418.5 7.0 3279 4 US-09-766-253-1
Sequence 1, Appli
89 403.5 6.8 5544 3 US-08-851-843A-68
Sequence 68, Appl
90 403.5 6.8 5544 3 US-08-974-549A-111
Sequence 111, App
91 403.5 6.8 5544 3 US-08-854-050-68
Sequence 68, Appl
92 403.5 6.8 5544 3 US-09-430-323-68
GENERAL INFORMA
93 403.5 6.8 5544 4 US-08-912-951-111
Sequence 111, App
94 403.5 6.8 5544 4 US-09-402-181B-111
GENERAL INFORMA
95 403.5 6.8 5544 4 US-09-721-456-111
GENERAL INFORMA
96 403.5 6.8 5544 4 US-09-766-253-68
Sequence 1, Appli
97 360 6.0 2714 4 US-09-417-485D-1
Sequence 3, Appli
98 356 6.0 2714 4 US-09-417-485D-3
Sequence 66, Appl
99 355.5 6.0 2631 3 US-08-851-843A-66
Sequence 225, App
100 355.5 6.0 2631 3 US-08-974-549A-225

ALIGNMENTS

RESULT 1

US-08-974-549A-343
; Sequence 343, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 343:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4037 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..3454
; OTHER INFORMATION: /note= "refined sequence of hTERT cDNA"
US-08-974-549A-343
Alignment Scores:
Pred. No.: 0 Length: 4037
Score: 5963.00 Matches: 1132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-08-974-584C-118 (1-1132) x US-08-974-549A-343 (1-4037)
QY 1 MetProArgAlaProArgCysArgAlaValArgSerLeuLeuArgSerHisTyrArgGlu 20
Db 56 ATGCGGGCGCTCCCGCTGCCGAGCGGTGCGCTCCCTGCTGCGCAGCACTACCGCGAG 115
QY 21 ValLeuProLeuAlaThrPheValArgArgLeuGlyProGlnGlyTyrArgLeuValGln 40
Db 116 GTGCTGCGCTGGCGCACGTTCTGTCGCGCGCTGTCGCGGCCAGCGCTGCTGTCGAG 175
QY 41 ArgGlyAspProAlaAlaPheArgAlaLeuValAlaGlnCysLeuValCysValProTrp 60
Db 176 GCGGGGACCGCGCGGGCTTTCCGCGCGCTGTCGCGGCCAGTCCCTGCTGCTGCGCCCTGG 235
QY 61 AspAlaArgProProAlaAlaProSerPheArgGlnValSerCysLeuLysGluLeu 80
Db 236 GACGACGGCG 295
QY 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaLysAenValLeuAlaPheGly 100
Db 296 GTGGCCCGAGTGTCTGACAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355
QY 101 PheAlaLeuLeuAspGlyAlaArgGlyClyProProGluAlaPheThrSerValArg 120
Db 356 TTCGCGCTGTGGACGGCG 415
QY 121 SerTyrLeuProAsnThrValThrAspAlaLeuArgGlySerGlyAlaTrpGlyLeuLeu 140
Db 416 AGCTACCTGCGCCCAACACGGTGACCGACATGCGGGGGAGCGGGGGCGTGGGGGGCTGTG 475
QY 141 LeuArgArgValGlyAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
Db 476 CTGCGCGCGGTGGCGACGACGCTGCTGTTTTCCTGCTGCGCACGCTGCTGCTGCTGCTG 535
QY 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
Db 536 CTGTGTGCTCCAGCTGCGCTACAGGTGTGCGGGCGCGCGCGCGCGCGCGCGCGCGCG 595
QY 181 AlaThrGlnAlaArgProProHisAlaSerGlyProArgArgArgLeuGlyCysGlu 200

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:24:56 ; Search time 912 Seconds
(without alignments)
7347.749 Million cell updates/sec

Title: US-08-974-584c-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRSLLRSYRE.....TALEAANPALPSPDKTILD 1132

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US08974584/runat.25072005.103112.9051/app_query.fasta_1.1287
-DB=N Geneseq_16Dec04 -QFWT=fastap -SUFFIX=n2p.rng -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pf0 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US08974584@cgn.1.1.455 @runat.25072005.103112.9051 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5963	100.0	3453	8	ABZ76217 Human TER
2	5963	100.0	4037	2	AAV22428 Human TER
3	5963	100.0	13766	6	AAZ46790 PGRN145 p
4	5952	99.8	3396	4	AAH48235 Heart mus
5	5952	99.8	3396	4	AAH49601 Human cat

6	5952	99.8	3396	4	AAH44366
7	5952	99.8	3396	12	ADG70113
8	5952	99.8	3399	6	ABV78144
9	5952	99.8	3399	6	ABZ35720
10	5952	99.8	3399	6	ABX09963
11	5952	99.8	3399	6	ABL91685
12	5952	99.8	3399	12	ADG90598
13	5952	99.8	3798	2	AAV27876
14	5952	99.8	3955	2	AAV22379
15	5952	99.8	4015	2	AAZ00724
16	5952	99.8	4015	2	AAZ20279
17	5952	99.8	4015	2	AAZ30154
18	5952	99.8	4015	4	AAH45901
19	5952	99.8	4015	6	AAZ46821
20	5952	99.8	4015	6	ABA97534
21	5952	99.8	4015	8	ACC57552
22	5952	99.8	4015	8	ABZ22474
23	5952	99.8	4015	8	ACC44482
24	5952	99.8	4015	8	ABZ18391
25	5952	99.8	4015	10	ACC58039
26	5952	99.8	4015	12	ADG85223
27	5952	99.8	4015	12	ADI82171
28	5952	99.8	4015	13	ADR70481
29	5952	99.8	4042	2	AAV72117
30	5952	99.8	4070	6	ABL53711
31	5947	99.7	4015	2	AAZ08150
32	5945	99.7	3396	2	AAZ18266
33	5945	99.7	3964	2	AAZ18254
34	5945	99.7	4023	2	AAV60320
35	5945	99.7	4027	2	AAZ89424
36	5945	99.7	4027	3	AAZ29388
37	5945	99.7	4027	10	ADC47060
38	5945	99.7	4027	10	ADE40481
39	5936	99.5	8742	6	AAZ46793
40	5994	98.8	3500	2	AAZ18275
41	5873	98.5	3918	2	AAZ18269
42	5873	98.5	3918	2	AAZ18278
43	5862	98.3	3543	8	ABZ69628
44	5822	94.3	3203	2	AAZ18268
45	5592	93.8	3323	2	AAZ18277
46	5584.5	93.7	3855	2	AAV22382
47	5543	93.0	3167	2	AAZ18271
48	5543	93.0	3167	2	AAZ18280
49	5432	91.1	7688	2	AAZ18351
50	5272	88.4	7797	2	AAZ18350
51	5002	83.9	3069	2	AAZ18267
52	4998	83.8	3069	2	AAZ18276
53	4926	82.6	3033	2	AAZ18270
54	4926	82.6	3033	2	AAZ18279
55	4901.5	82.2	7615	2	AAZ18349
56	4894	82.1	2848	2	AAV27872
57	4421.5	74.1	2541	2	AAZ18265
58	4421.5	74.1	2541	2	AAZ18274
59	3619	60.7	3387	12	ADG90602
60	3496	58.6	3369	12	ADG90600
61	3378	56.6	3972	2	AAV72125
62	3318	55.6	1883	2	AAZ18264
63	3154	52.9	2031	2	AAZ18263
64	3147	52.8	2031	2	AAZ18272
65	3125	52.4	1866	2	AAZ88251
66	3038	50.9	2357	2	AAZ15923
67	2789	46.8	51552	6	AAZ96607
68	2782	46.7	4335	2	AAV16979
69	2782	46.7	15418	3	AAZ63785
70	2782	46.7	15418	3	AAZ18601
71	2782	46.7	15418	6	ABZ54997
72	2782	46.7	15418	6	AAZ18595
73	2782	46.7	15418	10	ADC21253
74	2718	45.6	3346	2	AAV60321
75	2670.5	44.8	2175	2	AAZ15925
76	2661.5	44.6	2176	2	AAV22380
77	2582.5	43.3	2171	2	AAV22426
78	2451	41.1	2089	2	AAV72126

79 2178 36.5 1314 2 AAX88250
 80 2163.5 36.3 1311 2 AAX88243
 81 2159.5 36.2 1231 12 ADG90606
 82 2007 33.7 1728 12 ADG90604
 83 1868 31.3 1043 2 AAX18273
 84 1491 25.0 1012 2 AAV72124
 85 1484 24.9 873 9 AAL60416
 86 1456 24.4 2855 12 ADG90608
 87 1356.5 22.7 7498 3 AAX63786
 88 1356.5 22.7 7498 6 ABL50133
 89 1208.5 20.3 1096 2 AAX80994
 90 1052 17.6 949 2 AAV27875
 91 902 15.1 523 2 AAX15926
 92 898 15.1 519 9 AAL60417
 93 645 10.8 395 2 AAX15924
 94 565 9.5 2448 12 ADG70139
 95 555.5 9.3 2409 12 ADG70138
 96 555 9.3 1758 12 ADG70136
 97 478 8.0 1728 12 ADG70111
 98 461 7.7 2448 12 ADG70137
 99 453 7.6 2651 2 AAX80996
 100 443.5 7.4 734 4 AAK61529

ALIGNMENTS

RESULT 1
 ABZ76217 standard; DNA; 3453 BP.
 XX

AC ABZ76217;

DT 12-JUN-2003 (first entry)

XX Human TERT coding DNA fragment.

XX TERT; TRT; bone marrow stromal cell; MSC; telomeric repeat subunit;
 KW osteopathic; vulnary; gene therapy; human; ds.

OS Homo sapiens.

XX WO2003010305-A1.

XX 06-FEB-2003.

XX 26-JUL-2002; 2002WO-DK000514.

XX 27-JUL-2001; 2001DK-00001148.

XX (ARHU-) ARHUS AMT.

XX (UYAR-) UNIV ARHUS.

XX Thomas GJ, Moustapha K, Suresh ISR;

XX WPI; 2003-248078/24.

XX Immortalizing human stem cells for treating osteoporosis by culturing
 PT human bone marrow stromal cells, transducing the cultures with a
 PT retroviral vector, and obtaining an immortal stem cell line.

PS Disclosure; Fig 3-5; 38pp; English.

XX The invention relates to immortalizing human stem cells. The method
 CC involves (a) culturing human bone marrow stromal cells (hMSC); (b)
 CC transducing the cultures with a retroviral vector, comprising the human
 CC telomeric repeat subunit (htrt) gene; and (c) obtaining an immortal stem
 CC cell line. The immortalized stem cell line has substantially identical
 CC characteristics and properties when compared to the bone marrow stromal
 CC cells. The method is useful for metabolic sink (lowering blood
 CC cholesterol level) treating bone loss associated with ageing, bone
 CC fracture and/or osteoporosis or for tissue engineering e.g., creating
 CC bone or cartilage matrix for treating bone or cartilage defects and non-
 CC healed fractures, creating skin matrix for treating skin defects or

CC burns, or producing and secreting growth factors, such as VEGF, PDGF or
 CC hGH. The immortalized stem cells are useful for drug delivery of
 CC compounds, such as Coagulation factor VIII, Coagulation factor IX,
 CC Erythropoietin, insulin, leptin, angiotensins/endostatsins, human growth
 CC hormone and/or interleukins, for drug testing (identifying drugs acting
 CC on mesenchymal stem cell and/or drugs enhancing osteoblasts
 CC differentiation and/or drugs enhancing chondrocytes differentiation, for
 CC gene therapy or for producing high amounts of pure protein for
 CC crystallization. The present sequence represents a human TERT coding
 CC fragment, which was incorporated into a retroviral vector

XX SQ Sequence 3453 BP; 558 A; 1176 C; 1107 G; 612 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 8.33e-218 Length: 3453
 Score: 5963.00 Matches: 1132
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0

US-08-974-584C-118 (1-1132) x ABZ76217 (1-3453)

Qy 1 MetProArgAlaProArgCysArgAlaValArgSerLeuLeuArgSerHisTyArgGlu 20
 Db 10 ATGCCGCGCGCTCCCGCTGCCGAGCGCTCCCTCTGCTGCGCAGCCACTACCGCGAG 69
 Qy 21 ValLeuProLeuAlaThrPheValArgArgLeuGlyProGlnGlyTrpArgLeuValGln 40
 Db 70 GTGCTGCGCTGGCCACGTTGCTGCGCGCGCTGCGGCGCCCGAGGCTGGCGCTGGTGAG 129
 Qy 41 ArgGlyAspProAlaAlaPheArgAlaLeuValAlaGlnCysLeuValCysValProTrp 60
 Db 130 CGCGGGACCCCGCGGGCTTTCCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGCTGG 189
 Qy 61 AspAlaArgProProProAlaAlaProSerPheArgGlnValSerCysLeuLysGluLeu 80
 Db 190 GAGCACG 249
 Qy 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaLysAsnValLeuAlaPheGly 100
 Db 250 GTGGCCGCGAGTGTGCGAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 309
 Qy 101 PheAlaLeuLeuAspGlyAlaArgGlyGlyProGluAlaPheThrThrSerValArg 120
 Db 310 TTGCGCGTGTGCGACGCGGGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 369
 Qy 121 SerTyrLeuProAsnThrValThrAspAlaLeuArgGlySerGlyAlaTrpGlyLeuLeu 140
 Db 370 AGCTACCTGCCCAACACGCTGACCGACACTGCGGGGAGCGGGGCGTGGGGCTGCTG 429
 Qy 141 LeuArgArgValGlyAspAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
 Db 430 CTGCGCGCGTGGCGGCGAGCGACGTGCTTCCCTGCTGGCACGCTGCGCGCTCTTTGT 489
 Qy 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
 Db 490 CTGTGTGCTCCAGCTGCGCTACAGGTGTGCGGGCGCGCGCGCGCGCGCGCGCGCT 549
 Qy 181 AlaThrGlnAlaArgProProProHisAlaSerGlyProArgArgArgLeuGlyCysGlu 200
 Db 550 GCCACTCAGGCG 609
 Qy 201 ArgAlaTrpAsnHisSerValArgGluAlaGlyValProLeuGlyLeuProAlaProGly 220
 Db 610 CGGGCGCTGGACCATAGCGTCAGGAGCGCGGGGTCCCTCGGGCGCTGCCAGCCCCGGT 669
 Qy 221 AlaArgArgGlyGlySerAlaSerArgSerLeuProLeuProLysArgProArgArg 240
 Db 670 GCGAGAGGCGCGGGGCGAGTGCAGCGAAGTCTGCGCTTCCCGAAGAGGCCCGAGGCT 729
 Qy 241 GlyAlaAlaProGluProGluArgThrProValGlyGlnGlySerTrpAlaHisProGly 260

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 26, 2005, 11:25:31 ; Search time 7769 Seconds

(without alignments)
7060.279 Million cell updates/sec

Title: US-08-974-584C-118

Perfect score: 5963

Sequence: 1 MPRAPRCRAVRSLLRSHYRE.....TALEAANPALPSPFKTILD 1132

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q/cn2_1/USPTO spool/US08974584/runat_25072005_103112_9061/app_query.fasta_1.1287
-DB=genEmbl -QFMT=fastap -SURFIX=n2p.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biolum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US08974584 @CN 1.1 4648 @runat_25072005_103112_9061 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	5963	100.0	4037	6	E36819 Human telom
2	5963	100.0	4037	6	AR390496 Sequence
3	5963	100.0	4037	6	AR393110 Sequence
4	5963	100.0	4037	6	AX810378 Sequence

5	5963	100.0	4037	6	BD011070	Human tel
6	5963	100.0	13766	6	AX553919	Sequence
7	5952	99.8	3396	6	AR393358	Sequence
8	5952	99.8	3396	6	AR393359	Sequence
9	5952	99.8	3396	6	AR393360	Sequence
10	5952	99.8	3396	6	AR393361	Sequence
11	5952	99.8	3396	6	AR393362	Sequence
12	5952	99.8	3396	6	AX957656	Sequence
13	5952	99.8	3396	6	BD091553	Adult bon
14	5952	99.8	3396	6	BD094749	The cell
15	5952	99.8	3396	6	BD096291	Cells cap
16	5952	99.8	3399	6	AX481414	Sequence
17	5952	99.8	3451	6	AR393441	Sequence
18	5952	99.8	4015	6	AR104587	Sequence
19	5952	99.8	4015	6	AR175848	Sequence
20	5952	99.8	4015	6	BD218834	Method an
21	5952	99.8	4015	6	E36793	E36793 Human telom
22	5952	99.8	4015	6	AR182221	Sequence
23	5952	99.8	4015	6	AR224455	Sequence
24	5952	99.8	4015	6	AR226390	Sequence
25	5952	99.8	4015	6	AR243328	Sequence
26	5952	99.8	4015	6	AR263555	Sequence
27	5952	99.8	4015	6	AR265996	Sequence
28	5952	99.8	4015	6	AR390470	Sequence
29	5952	99.8	4015	6	AR393084	Sequence
30	5952	99.8	4015	6	AR404030	Sequence
31	5952	99.8	4015	6	AR438403	Sequence
32	5952	99.8	4015	6	AX019310	Sequence
33	5952	99.8	4015	6	AX133979	Sequence
34	5952	99.8	4015	6	AX552695	Sequence
35	5952	99.8	4015	6	AX810036	Sequence
36	5952	99.8	4015	6	BD011044	Human tel
37	5952	99.8	4015	6	BD015832	Quantitat
38	5952	99.8	4015	6	BD082985	Method fo
39	5952	99.8	4015	6	BD131727	Method fo
40	5952	99.8	4015	9	AF015950	Homo sapi
41	5952	99.8	4042	6	BD136185	Human tel
42	5952	99.8	4042	6	AX001446	Sequence
43	5952	99.8	4042	6	AX003121	Sequence
44	5952	99.8	4070	6	AX391846	Sequence
45	5950	99.8	4015	6	C0717133	Sequence
46	5945	99.7	3396	6	BD196290	Vertebrat
47	5945	99.7	3964	6	BD196267	Vertebrat
48	5945	99.7	4027	6	BD233924	Method an
49	5945	99.7	8960	12	AF043739	Synthetic
50	5945	99.5	8742	6	AX553924	Sequence
51	5936	99.5	3500	6	BD196299	Vertebrat
52	5894	98.8	3544	6	AX752636	Sequence
53	5894	98.8	3918	6	BD196293	Vertebrat
54	5873	98.5	4022	6	BD196302	Vertebrat
55	5822	97.6	4022	6	BD196292	Vertebrat
56	5585.5	93.7	3362	6	E36795	Human telom
57	5584.5	93.7	3855	6	AR243330	Sequence
58	5584.5	93.7	3855	6	AR390472	Sequence
59	5584.5	93.7	3855	6	AR393086	Sequence
60	5584.5	93.7	3855	6	AX810039	Sequence
61	5584.5	93.7	3855	6	BD011046	Human tel
62	5584.5	93.5	2210	9	AB085628	Homo sapi
63	5574.5	93.4	4029	6	AR104586	Sequence
64	5570	93.4	4029	6	AR175847	Sequence
65	5570	93.4	4029	6	E36818	Human telom
66	5570	93.4	4029	6	AR390495	Sequence
67	5570	93.4	4029	6	AR393109	Sequence
68	5570	93.4	4029	6	AX810327	Sequence
69	5570	93.4	4029	6	BD011069	Human tel
70	5570	93.4	4029	6	BD196301	Vertebrat
71	5534.5	92.8	3466	6	BD196295	Vertebrat
72	5506.5	92.3	3326	6	BD196304	Vertebrat
73	5460.5	91.6	3432	6	BD196307	Vertebrat
74	5432	91.1	7688	6	BD196306	Vertebrat
75	5272	88.4	7797	6	BD196306	Vertebrat
76	5207	87.3	3028	9	AB086379	Homo sapi
77	5128	86.0	2992	9	AB086950	Homo sapi

c

```

78 5005 83.9 3069 6 BD196291 Vertebrat
79 4954 83.1 3173 6 BD196300 Vertebrat
80 4926 82.6 3033 6 BD196294 Vertebrat
81 4904.5 82.2 7615 6 BD196305 Vertebrat
82 4875 81.8 3137 6 BD196303 Vertebrat
83 4421.5 74.1 2541 6 BD196289 Vertebrat
84 4370.5 73.3 2645 6 BD196298 Vertebrat
85 4081 68.4 3419 4 AF380351 Canis fam
86 3619 60.7 4170 10 AF149012 Mesocric
87 3520.5 59.0 3378 10 AY5339717 Rattus no
88 3506 58.8 2041 6 BD196288 Vertebrat
89 3505.5 58.8 3360 10 AY5339718 Rattus no
90 3496 58.6 3369 10 AF073311 Mus muscu
91 3496 58.6 3426 10 AF051911 Mus muscu
92 3496 58.6 3496 6 AR565399 Sequence
93 3455 57.9 2145 6 BD196297 Vertebrat
94 3419 57.3 3483 10 AY5339719 Rattus no
95 3378 56.6 3972 6 BD136189 Human tel
96 3378 56.6 3972 6 AX001451 Sequence
97 3360.5 56.4 3451 10 AY5339720 Rattus no
98 3233 54.2 3747 10 BC082327 Mus muscu
99 3154 52.9 2031 6 BD196287 Vertebrat
100 3125 52.4 1866 6 E27468 Novel gene

```

ALIGNMENTS

```

RESULT 1
E36819 4037 bp DNA linear PAT 18-JUN-2001
LOCUS Human telomerase catalytic subunit promoter.
DEFINITION Human telomerase catalytic subunit promoter.
ACCESSION E36819
VERSION E36819.1 GI:13022782
KEYWORDS JP 1999253177-A/27.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4037)
AUTHORS Thomas,R.S., Jochimu,R., Toru,N., Karen,B.C., Greg,B.M.,
Calvin,B.H. and William,H.A.
TITLE Human telomerase catalytic subunit promoter
JOURNAL Patent: JP 1999253177-A 27 21-SEP-1999;
JERON CORP, UNIVERSITY TECHNOLOGY CORP
COMMENT OS Unidentified
PN JP 1999253177-A/27
PD 21-SEP-1999
PF 15-OCT-1998 JP 1998320169
PR 01-OCT-1996 US 08/724,643,18-APR-1997 US 08/844,419, PR
25-APR-1997 US 08/846,017,06-MAY-1997 US 08/851,843, PR
09-MAY-1997 US 08/854,050,14-AUG-1997 US 08/911,312, PR
14-AUG-1997 US 08/912,951,14-AUG-1997 US 08/915,503 PI THOMAS
R SECHI, JOCHIMU RINGER, TORU NAKAMURA, KAREN B CHAPMAN, PI GREG B
MORIN,
PI CALVIN B HAREI, WILLIAM H ANDREWS
PC C12N15/09, A61K31/70, A61K38/55, A61K39/395, A61K48/00,
PC C12Q1/02
PC C12Q1/48, C12Q1/68, G01N33/15, G01N33/48, G01N33/50//C07K14/47, PC
C07K16/40,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/12, C12P21/08, (C12N1/19, PC
C12R1:84),
PC (C12N9/12, C12R1:19), (C12N9/12, C12R1:19), (C12N9/12, C12R1:84),
PC (C12N9/12, C12R1:91), C12N15/00, A61K37/64, C12N5/00 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS
1..4037
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:3264"

```

FEATURES
source

ORIGIN

```

Alignment Scores:
Pred. No.: 3.66e-150 Length: 4037
Score: 5963.00 Matches: 1132
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-08-974-584C-118 (1-1132) x E36819 (1-4037)
Qy 1 MetProArgAlaProArgCysArgAlaValArgSerLeuArgSerHisTyrArgGlu 20
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGGTGCGCTCCCTGTGTCGCGAGCCACTACCGCGAG 115
Qy 21 ValLeuProLeuAlaThrPheValArgLeuGlyProGlnGlyTTPArgLeuValGln 40
Db 116 GTGCTGCGCTGGCCAGTTCGTGCGCGCTGGGGGCCCGAGGCTGGCGCTGGTGAG 175
Qy 41 ArgGlyAspProAlaAlaPheArgAlaLeuValAlaGlnCysLeuValCysValProTrp 60
Db 176 CGCGGGACCGCGCGGTTCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGTGG 235
Qy 61 AspAlaArgProProProAlaAlaProSerPheArgGlnValSerCysLeuGlyGluLeu 80
Db 236 GACGACGCGCGCGCGCGCTCCCTTCGCGCAGGTGCTGCTGCTGAGGAGCTG 295
Qy 81 ValAlaArgValLeuGlnArgLeuCysGluArgGlyAlaValAenValLeuAlaPheGly 100
Db 296 GTGCGCGCGAGTGTGCGAGAGCTGTGCGAGCGCGCGCGAGAACGTGCTGGCTTCGGC 355
Qy 101 PheAlaLeuLeuAspGlyAlaArgGlyGlyProProGluAlaPheThrThrSerValArg 120
Db 356 TTGCGCGTGTGAGCGGGCGCGCGGGCGCGCGCGCGCGCGCGCTTACCACCGAGTGGC 415
Qy 121 SerTyrLeuProAenThrValThrAspAlaLeuArgGlySerGlyAlaTTPArgLeuLeu 140
Db 416 AGCTACCTGCGCAACAGCGTGACGAGCAGCTGCGGGGAGCGGGGCTGGGGCTGCTG 475
Qy 141 LeuArgArgValGlyAspAspValLeuValHisLeuLeuAlaArgCysAlaLeuPheVal 160
Db 476 CTGCGCGCGTGGCGAGCAGCTGCTGGTTACCTGTGCGCGCGCGCGCTCTTTGTG 535
Qy 161 LeuValAlaProSerCysAlaTyrGlnValCysGlyProProLeuTyrGlnLeuGlyAla 180
Db 536 CTGCTGGCTCCAGCTGCGCTACAGGTGTGCGGGCGCGCGCTGTACCAGCTCGGCGCT 595
Qy 181 AlaThrGlnAlaArgProProProHisAlaSerGlyProArgArgLeuGlyCysGlu 200
Db 596 GCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGGGATGCGAA 655
Qy 201 ArgAlaTrpAenHisSerValArgGluAlaGlyValProLeuGlyLeuProAlaProGly 220
Db 656 CGGCGCTCGGAACCATAGCGTACGAGGCGCGGGGTCCCTTGGGCGCTGGCAGCCCCGGGT 715
Qy 221 AlaArgArgArgGlyGlySerAlaSerArgSerLeuProLeuProLysArgProArgArg 240
Db 716 GCGAGGAGGCGCGGGGCGAGTGCAGCGAGTCTGCGCTGCCAGAGGCGCGCGCGCT 775
Qy 241 GlyAlaAlaProGluProGluArgThrProValGlyGlnGlySerTrpAlaHisProGly 260
Db 776 GCGCTGCGCTGAGCGGAGCGGAGCGCGCGCTTGGGCGAGGGGTCTGCGGCGCGCGCG 835
Qy 261 ArgThrArgGlyProSerAspArgGlyPheCysValValSerProAlaArgProAlaGlu 280
Db 836 AGGACGCGTGGACCGAGTACCGGTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
Qy 281 GluAlaThrSerLeuGluGlyAlaLeuSerGlyThrArgHisSerHisProSerValGly 300
Db 896 GAACCGACCTCTTTGGAGGGTGGCTCTCTGGGACGCGGCACCTCCACCCCATCCGTGG 955
Qy 301 ArgGlnHisAlaGlyProProSerThrSerArgProArgProArgProTrpAspThrPro 320
Db 956 CGCCAGCACACGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1015

```